INVESTIGATING THE MOLECULAR BASIS OF JEUNE AND CENANI-LENZ SYNDROMES

Chiara Bacchelli

A thesis submitted for the degree of Doctor of Philosophy
to the University of London

Molecular Medicine Unit
Institute of Child Health, London

April 2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

This thesis describes an investigation into the molecular basis of two rare autosomal recessive conditions: Jeune asphyxiating thoracic dystrophy (JATD) and Cenani-Lenz syndrome (CLS).

JATD is a skeletal dysplasia sometimes complicated by chronic nephritis, hepatic and pancreatic fibrosis, and retinal degeneration. Although it is a serious condition with a high mortality and morbidity, predominantly from respiratory insufficiency in infancy and chronic renal failure in childhood, its molecular basis is currently unknown. This project focused on twelve consanguineous and nineteen non-consanguineous affected families. Haplotype analysis was undertaken in the consanguineous families to exclude linkage to several previously proposed candidate genes and loci. A homozygosity mapping approach was then adopted, and a genome-wide linkage scan was carried out in the three most statistically powerful consanguineous families. Each family was consistent with linkage to a different locus, on chromosomes 3q24-q26, 8q24 and 14q24-q32. The remaining families were genotyped across each locus to define minimal critical intervals. A number of candidate genes were then selected and screened for mutations in families which were consistent with linkage, but no mutations were identified.

CLS is a distal limb malformation characterised by total digit syndactyly and radio-ulnar synostosis. *Limb deformity* (*ld*) mice have a very similar phenotype. The *ld* phenotype is caused by mutations in the *Formin-1* gene, leading to loss of *Gremlin* expression in the limb bud mesenchyme. In this project, haplotype analysis in one consanguineous CLS family excluded linkage both to *FORMIN-1* (*FMN-1*) and to *GREMLIN*, which are immediately adjacent to each other on chromosome 15q13. Haplotype analysis in another consanguineous CLS family was consistent with linkage to this region, but no mutations were identified in either *FMN-1* or *GREMLIN*. The coding sequence of *GREMLIN* was analysed in two further consanguineous and three non-consanguineous CLS families, but no mutations were identified.

List of contents

List of figures

List of tables

List of appendices

Aknowledgements

Firstly, I would like to thank my supervisors, Dr Frances Goodman and Prof. Peter Scambler for their encouragement and support during the course of this project and all my colleagues, past and present, in the Molecular Medicine Unit for their precious help. Special thanks go to Frances, who introduced me to PCR, sequencing and lab work in general and Neil, who gave me important tips on genotyping and homozygosity mapping and has been very supportive throughout this project. A big thank you also goes to my office-companions Beth, Katrina, Lesley, Paris and Phil for their encouragements and laughs.

I would also like to thank Dr Louise Wilson for her advice and assistance with the clinical aspects of the project and allowing me to sit in during her Clinic and meet one of the JATD families. I am also grateful to all the clinicians who have helped to recruit families and who have sent us DNA samples and our collaborators Dr Colin Johnson and Neil Morgan in Birmingham and Dr Valerie Cormier-Daire in Paris. A particular thanks go to Dr Christine Hall for her help with the interpretation of radiographs.

I must pay particular thanks to my family in Italy, la mia mamma, il mio papá, Anna (and especially my beloved Grandmother, who always believed in me and in my ability to succeed and who I do miss immensely) and all my friends in Italy for being always by my side, even if miles away. Grazie ragazzi! A very special thanks also to my very best friend Leila here in London- I could not have done without her.

But above all, my very special thanks are for Paolo, who has been my real strength over the years. He has shared with me all the frustrations of thousands of PCRs and put up with the mad writing up period, but also shared the joy and excitement that came with that.

Finally, thanks to Andrea, always by my side during the thesis writing, for teaching me what is really important in life, and to never give up ("va tutto bene"-"it's all good" he used to say). I do miss you a lot...

List of abbreviations

A	adenine
aa	amino acid
AER	apical ectodermal ridge
APS	ammonium persulphate
BAC	bacterial artificial chromosome
BBS	Bardet-Biedl syndrome
BLAST	basic local alignment search tool
BMP	bone morphogenetic protein
bp	base pair(s)
C	cytosine
°C	degree Celsius
cDNA	complementary DNA
cM	centiMorgan
CLS	Cenani-Lenz syndrome
DMSO	dimethyl sulphoxide
dNTP	deoxynucleoside triphosphate
ddNTP	dideoxynucleotide
DNA	deoxyribonucleic acid
EDTA	ethylenediamine tetra-acetic acid
ESE	exonic splicing enhancer
EVC	Ellis-van Creveld
FBP	formin binding protein
FBS	fetal bovine serum
Fgf	fibroblast growth factor
FH	formin homology
FMN-1	*FORMIN-1*
G	guanine
HFGS	Hand-foot-genital syndrome
IBD	identity by descent

IBS	identity by state
JATD	Jeune asphyxiating thoracic dystrophy
kb	kilobase pair
KO	knockout
l	litre
ld	*limb deformity*
LDE	linkage disequilibrium
lod	logarithm of the odds
M	Molar
mM	millimoles per litre
Mb	megabase pair
μ	micro
m	milli
mRNA	messenger RNA
n	nano
p	pico
p	short arm of chromosome
PACs	P1 artificial chromosomes
PCR	polymerase chain reaction
PPD	preaxial polydactyly
q	long arm of chromosome
RNA	ribonucleic acid
SD1	syndactyly type 1
Shh	*Sonic hedgehog*
SHOX2	*Short stature homeobox 2*
SNP	single nucleotide polymorphism
SRPS	short rib-polydactyly syndromes
srt	*shorty*
θ	recombination fraction
T	thymine
TBE	Tris-borate/ EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
TGFβ	transforming growth factor β family

TPRs	tetratricopeptide repeats
TTC7B	Tetratricopeptide repeat domain 7B
US	ultrasound
UTR	untranslated region
UV	ulteviolet
V	volt
WISP1	*Wnt1-inducible signalling pathway protein 1*
YACs	yeast artificial chromosomes
Y2H	Yeast two hybrid
ZPA	zone of polarizing activity

Chapter 1 Introduction: Jeune and Cenani-Lenz syndromes

1.1 Jeune syndrome

1.1.1 Clinical features

Jeune syndrome or asphyxiating thoracic dystrophy (JATD; MIM 208500) is a rare autosomal recessive skeletal dysplasia with renal, pancreatic, hepatic and retinal involvement. The disorder was first described in an affected sister and brother in 1954 and 1955 by Jeune (Jeune M, 1955; Jeune M, 1954). Over one hundred cases have been so far reported (Tongsong *et al.*, 1999). JATD has an incidence of 1 in 100,000-130,000 live births (Oberklaid *et al.*, 1977).

The main skeletal features of JATD consist of a small chest, mild short stature, mild shortening of limbs (usually rhizomelic) and post-axial polydactyly (in about half of cases). No developmental delay or facial dysmorphism are described. The thoracic malformation with short ribs usually leads to pulmonary hypoplasia with a marked reduction in alveoli, resulting in respiratory problems during the neonatal period (Finegold *et al.*, 1971; Langer, 1968; Turkel *et al.*, 1985). Up to 70% of patients die of respiratory failure and infections in early childhood, usually within the first year of life (Ozcay *et al.*, 2001). Nevertheless, advances in respiratory neonatal intensive care have now improved chances of survival (Kajantie *et al.*, 2001). In these cases, there is usually progressive improvement in the relative growth of the chest. There are characteristic skeletal abnormalities on radiological examination, which are essential to confirm the diagnosis. These include small bell-shaped thorax with short horizontal ribs and 'handlebar' clavicles (Fig. 1.1A); small iliac wings with a 'trident' appearance of the acetabular roofs (Fig. 1.1B); flaring of metaphyses of long bones (Fig. 1.1C and D); and, in the hands, short broad middle phalanges, hypoplastic distal phalanges and cone-

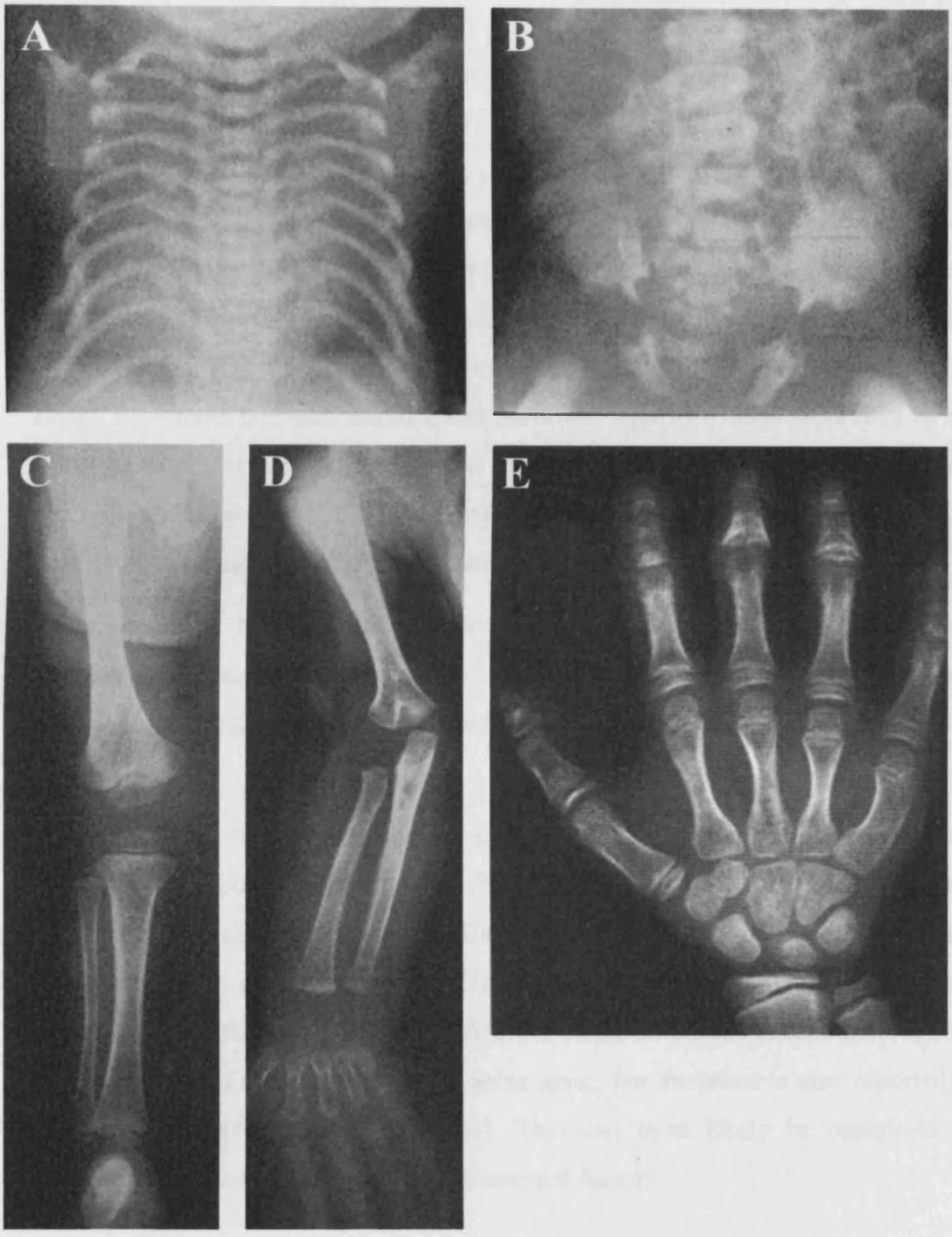


Figure 1.1 Radiographs showing the typical features of JATD. The thoracic cage is small and narrow with short ribs (A). The acetabular roofs have a 'trident' configuration (B). The proximal bones of the limbs are shorter than the middle bones (C and D). The middle and distal phalanges of the fingers are abnormally small (E).

shaped epiphyses (Fig. 1.1E).

In those patients who survive respiratory failure, chronic renal failure is the commonest cause of death (Langer, 1968; Turkel *et al.*, 1985; Shah, 1980; Shokeir *et al.*, 1971; Oberklaid *et al.*, 1977). It is therefore important to monitor renal function, in order to intervene with dialysis or renal transplant. Clinical problems due to pancreatic insufficiency may be present in some cases, and may require pancreatic enzyme supplements (Hopper *et al.*, 1979; Turkel *et al.*, 1985). Histologically, cystic changes and fibrosis in the pancreas have been described (Georgiou-Theodoropoulos *et al.*, 1988). Prolonged neonatal jaundice can occur and in severe cases there may be hepatic fibrosis and biliary cirrhosis with portal hypertension (Hudgins L, 1990; Whitley *et al.*, 1987; Labrune *et al.*, 1999). In the eyes, progressive retinal degeneration can occur, leading in some cases to blindness (Casteels *et al.*, 2000; Wilson *et al.*, 1987).

Although JATD is autosomal recessive, there are occasional reports of a possible heterozygous phenotype in parents and siblings of affected children consisting of short stature, mild chest deformity and brachydactyly (Ozcay *et al.*, 2001; Shokeir *et al.*, 1971).

The name JATD is rather misleading, as asphyxia does not occur in all cases and there are numerous reports of patients with only mild or no respiratory problems (Giorgi *et al.*, 1990; Oberklaid *et al.*, 1977; Capilupi *et al.*, 1996). Also, the extra-skeletal complications do not always occur. Therefore there is a wide variation in phenotypic severity. Variation between different families could be due to genetic heterogeneity or to different types of mutations in the same gene, but variation is also reported within affected families (Ozcay *et al.*, 2001). This can most likely be explained by the modifying effects of other genes/environmental factors.

1.1.2 Overlap with Ellis-van Creveld syndrome

The abnormalities in the rib cage, limbs and pelvis of JATD patients are extremely similar to those that occur in another autosomal recessive skeletal dysplasia, Ellis-van Creveld syndrome (EVC; MIM 225500). There are, however, some significant

differences between the two conditions. In EVC syndrome, 50-60% of patients have a congenital heart defect, which is not a feature of JATD, and many EVC patients also have other abnormalities not seen in JATD, including urogenital tract abnormalities (epispadias, hypospadias, cryptorchidism), dysplastic nails and teeth, and partially cleft upper lip (Brueton *et al.*, 1990). On the other hand, neonatal respiratory problems, involvement of the kidney, pancreas, liver, and retinal degeneration are features of JATD, but do not occur in EVC. The best distinguishing feature is the polydactyly, that is a constant feature in the hands but rarely affects the feet in EVC. In JATD, polydactyly is an inconstant feature, but when present, usually affects both the hands and the feet (Brueton *et al.*, 1990). Nevertheless, the similarities between the two syndromes have led some authors to suggest that they form part of the same disease spectrum rather than being distinct conditions (Brueton *et al.*, 1990; Moudgil *et al.*, 1998).

1.1.3 Overlap with short rib-polydactyly syndromes

JATD also shows considerable clinical overlap with the short rib-polydactyly syndromes (SRPS), a group of recessively-inherited lethal skeletal dysplasias all characterised by a hypoplastic thorax, short ribs, short limbs, polydactyly and visceral abnormalities. These syndromes have been classified into four different types (type I - Saldino-Noonan, type II - Majewski, type III - Varma-Naumoff, and type IV - Beemer-Langer), but they have many clinical, radiological and histological similarities, and may well represent different variants of the same disease spectrum, with a common cause or at least a common pathogenic pathway (Franceschini *et al.*, 1995). Some patients have in fact been reported with clinical/radiological features of all four types (Martinez-Frias *et al.*, 1993; Sarafoglou *et al.*, 1999). Ho and colleagues reported a case of JATD and SRPS type III occurring in the same consanguineous Asian family. One first-cousin marriage produced two brothers with typical mild JATD; another first-cousin marriage produced a boy with lethal SRPS type III. This suggests the two conditions may be different forms of the same disorder with the intrafamilial variation in severity reflecting the modifying influence of genetic/environmental factors (Ho *et al.*, 2000).

1.1.4 Candidate genes and loci

1.1.4.1 Two candidate genes: EVC and EVC2

A possible candidate gene for JATD was thought to be the gene mutated in EVC syndrome because of the overlapping features between the two conditions. The EVC syndrome has been mapped to a critical region on chromosome 4p16, with no evidence of genetic heterogeneity (Howard *et al.*, 1997). A first underlying gene, named *EVC*, was cloned and shown to be mutated in a number of EVC patients (Ruiz-Perez *et al.*, 2000). It codes for a novel, widely-expressed protein of unknown function with no homology to known proteins. *In situ* hybridisation in human embryonic tissue showed *EVC* expression in developing bones, heart, kidney and lung. In bone, *EVC* was expressed in the developing vertebral bodies, ribs and both upper and lower limbs with higher expression in the distal limb compared to the proximal limb.

In a patient with EVC syndrome, two homozygous mutations were subsequently identified in a second gene named *EVC2,* which lies immediately 5' of *EVC* (Galdzicka *et al.*, 2002). Mutations in *EVC* and *EVC2*, leading to an identical phenotype were also reported by Ruiz-Perez and colleagues (Ruiz-Perez *et al.*, 2003). The authors identified mutations in *EVC2* in six consanguineous EVC pedigrees consistent with localization to the EVC region but with no mutations in *EVC*. *EVC* and *EVC2* are non-homologous genes with unknown function arranged in a head-to-head configuration probably sharing a single promoter or elements of overlapping promoters (Ruiz-Perez *et al.*, 2003). *EVC2* is expressed in heart, placenta, lung, liver, skeletal muscle, kidney and pancreas. A transcript is also found in lymphoblasts and chondrocytes (Galdzicka *et al.*, 2002). The bovine orthologue of *EVC2* named *LIMBIN* (*LBN*) is mutated in bovine chondrodysplastic dwarfism, an autosomal recessive disorder characterised by short limbs and joint abnormalities. The mouse orthologue is strongly expressed in the proliferative chondrocytes of long bone epiphysial growth plates, adult heart and kidney (Takeda *et al.*, 2002).

However, in one family with JATD and one with SRPS type III, linkage to the *EVC/EVC2* locus has clearly been excluded (Krakow *et al.*, 2000). Nevertheless, the pathogenesis of the two conditions could well be similar, and the JATD gene could have similar biological and/or developmental functions or even act in the same molecular pathway as the *EVC* and/or the *EVC2* gene.

In addition, for all four types of SRPS, the underlying gene(s) have not yet been mapped or cloned; one or more of them could be allelic to JATD.

1.1.4.2 A candidate locus: 12p11-p12

Nagai and colleagues have described a boy with a *de novo* deletion on the short arm of chromosome 12, at 12p11.21-p12.2, who had a small rib cage, a hypoplastic pelvis, short stature and radiological findings similar to those in JATD and EVC syndromes (Nagai *et al.*, 1995). They therefore proposed that the gene underlying JATD might be located on chromosome 12p11-12. This patient also had features not seen in JATD, including mild developmental delay, oligodontia, and hypoplastic hair and skin. The breakpoints of the deletion were later fine-mapped and the deletion was found to remove an approximately 15 cM region between markers D12S1682 and D12S345 (Bahring *et al.*, 1997).

1.1.4.3 A candidate mouse model: shorty

A possible mouse model for JATD has been identified in a screen of embryos derived from mice mutagenised with N-ethyl-N-nitrosourea, which generates single-nucleotide mutations (Herron *et al.*, 2002). One mutant, named 'shorty' (*srt*), had an extremely small chest, and died soon after birth from respiratory failure. However, the skeletal abnormalities in this mouse are not exactly the same as in JATD patients. They include major vertebral anomalies, resulting in a shortened body axis and severe lordosis of the lower cervical/upper thoracic spine, absence of the proximal portion of the ribs and clefting of the secondary palate. None of these features are described in JATD.

Furthermore, there is no report of histological changes in the kidney, liver or pancreas, which typically occur in JATD.

The *srt* phenotype is similar to that of the 'pudgy' mouse, which is caused by mutations in *Dll3* on mouse chromosome 7 (Kusumi *et al.*, 1998). In humans, mutations in *DLL3* cause spondylocostostal dysostosis, a rare autosomal recessive skeletal dysplasia (Bulman *et al.*, 2000). Linkage analysis in *srt* excluded chromosome 7, and instead mapped the locus to a 7 cM interval on mouse chromosome 17. A good candidate gene in this region was *Dll1*, but no mutations in it were identified, so the underlying gene is at present unknown. In humans, regions corresponding to *srt* critical interval lie on 6p21, 6q25-q27 and 16p13.3.

1.2 Cenani-Lenz syndrome

1.2.1 Clinical features

Cenani-Lenz syndrome (CLS; MIM 212780) is a rare recessively-inherited congenital limb malformation, first described by Cenani and Lenz in 1967 (Cenani and Lenz, 1967) in two brothers with pronounced radioulnar and metacarpal synostosis. The upper limbs are usually much more severely affected than the lower limbs. In the hands, there is syndactyly between all the digits, with central digits often sharing the same nail; digits are short, and some may be missing (Fig. 1.2B and D). Radiographs show malformation, fusion and absence of most of the phalanges, metacarpals and carpals (Fig. 1.2A and C), often accompanied by partial or complete radio-ulnar synostosis, resulting in limitation of pronation and supination. In the feet there is usually just soft tissue syndactyly between some of the toes, and tibio-fibular synostosis has never been observed. Recently, four reports described abnormalities not previously associated with CLS, namely rib and vertebral abnormalities (Seven *et al.*, 2000), renal hypoplasia (Bacchelli *et al.*, 2001), facial dysmorphism (Temtamy *et al.*, 2003), congenital cataract and duplication of the big toes (Percin and Percin, 2003).

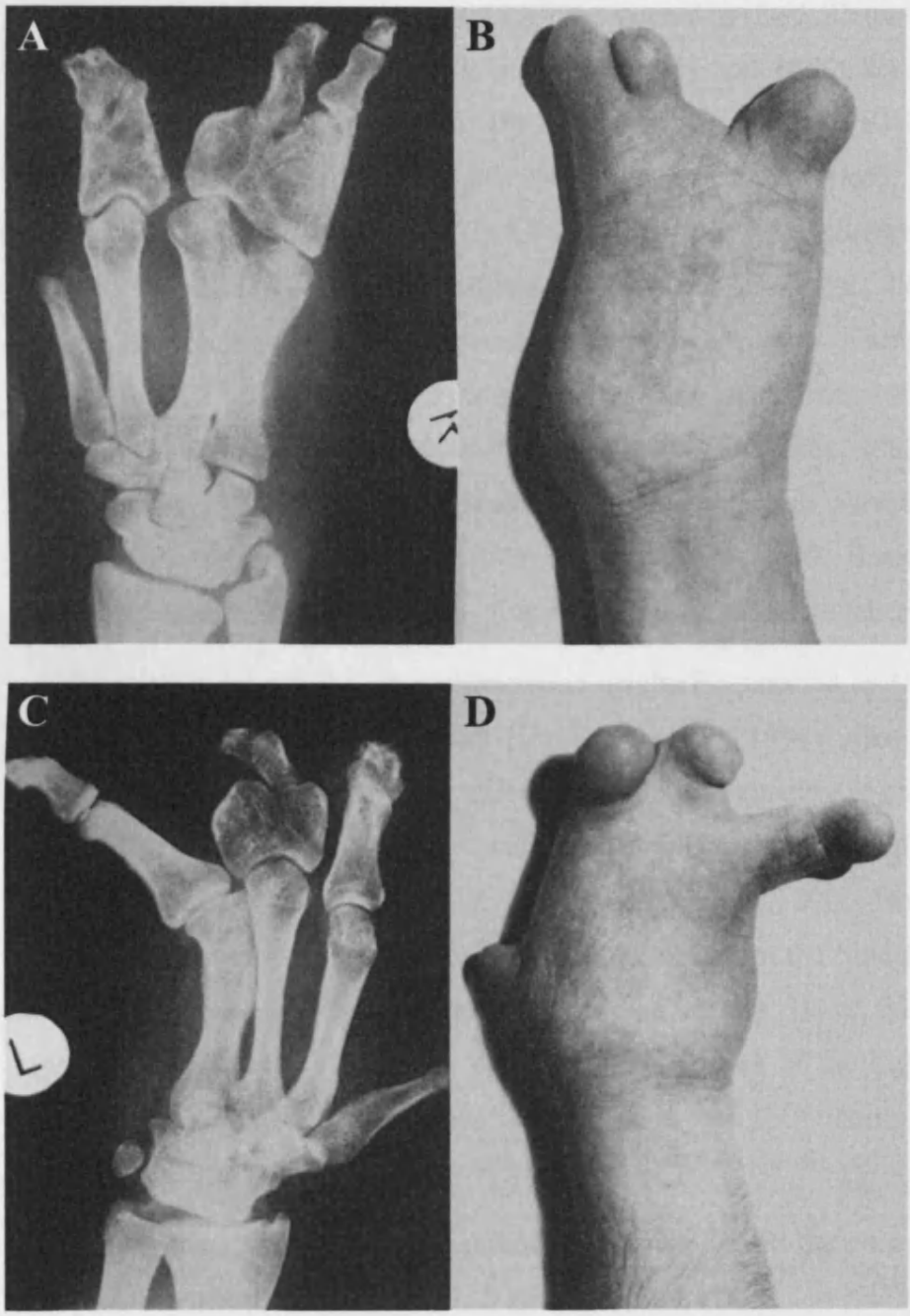


Figure 1.2. The hands of a patient with CLS. There is severe shortening and syndactyly of all the digits, with malformation, fusion or absence of many of the phalanges, metacarpals and carpals.

To date, only twenty-five cases of CLS have been described (Cenani and Lenz, 1967; Drohm *et al.*, 1976; Verma *et al.*, 1976; Temtamy and McKusick, 1978; Dodinval, 1979; Pfeiffer and Meisel-Stosiek, 1982; De Smet *et al.*, 1992; Elcioglu *et al.*, 1997; Seven *et al.*, 2000; Bacchelli *et al.*, 2001; Nezarati and McLeod, 2002; Percin and Percin, 2003; Temtamy *et al.*, 2003). Of these, two also have extreme shortening of the forearms, and may represent a separate condition (Drohm *et al.*, 1976; Verma *et al.*, 1976). Although most cases are sporadic, there are four reports of affected siblings born to unaffected non-consanguineous parents (Cenani and Lenz, 1967; Temtamy and McKusick, 1978; Dodinval, 1979; Pfeiffer and Meisel-Stosiek, 1982), and six reports of affected children born to unaffected consanguineous parents (Temtamy and McKusick, 1978; Verma *et al.*, 1976; Elcioglu *et al.*, 1997; Bacchelli *et al.*, 2001; Temtamy *et al.*, 2003), strongly suggesting autosomal recessive inheritance.

A dominantly-inherited limb malformation originally reported as CLS in an affected father and two affected daughters (De Smet *et al.*, 1996), and associated with a t(12;22)(p11.2;q13.3) balanced translocation, was subsequently re-classified as a novel type of complex synpolydactyly with metacarpal, metatarsal and tarsal synostosis (Debeer *et al.*, 1998a; Debeer *et al.*, 1998b; Debeer *et al.*, 2000; Debeer *et al.*, 2002). The chromosome 22 breakpoint has been shown to disrupt the fibulin-1 gene (*FBLN1*), which is expressed in the extracellular matrix in various tissues during development, including the interdigital regions and the perichondrium of the digits (Debeer *et al.*, 2002). Quasi-dominant inheritance is reported in one CLS family (Temtamy *et al.*, 2003).

The molecular basis of CLS is currently unknown and at the commencement of this project, no attempt to locate the CLS gene had been made.

1.2.2 Mouse mutants as models for human diseases

Mice are the most widely used animal models of human disease. Because of their small size, short lifespan and generation time, and because they can be bred easily and cheaply, large-scale breeding, genetic crosses and systematic mutagenesis programs

have been arranged and studied for decades (Wynshaw-Boris, 1996). A large number of mouse mutants have been generated following mutagenesis programs and phenotypes for many mutants have been described. Some mutants originated spontaneously, others have been generated artificially by a variety of different processes such as exposure to mutagenic chemicals, for example ethylnitrosourea (ENU), or high doses of X-rays. Chemically-induced and irradiation-induced mutagenesis however, has the disadvantage of being random. More recently, gene targeting, insertional mutagenesis and transgenic techniques have allowed the creation of animal disease models resembling corresponding specific human diseases by introducing mutations into a particular gene *in vivo* (Gordon and Ruddle, 1983). Mice can be generated with a particular alteration in a chosen target gene, or simply by disrupting a particular gene with a large insertional cassette, thereby generating null mutations (known as gene knockouts). Such techniques are largely used to model loss of function mutations.

Homologous mouse and human mutants often show similar phenotypes, offering valuable clues for identification of human disease genes. Almost all human genes have an easily identifiable mouse homolog because of the high level of conservation between human and mouse coding sequences. Because the homologous regions of mouse and human genomes have been well documented and become easily accessible (see recent Ensembl freeze at http://www.ensembl.org/ or human/mouse homology maps at http://www.ncbi.nlm.nih.gov/Omim/Homology) once a disease gene in one species has been identified, it is possible to predict the location of that gene in the other species. For example, the spontaneous mouse mutant *Hypodactyly* has similar limb abnormalities to those seen in the human condition hand-foot-genital syndrome (HFGS). When the *Hypodactyly* mouse was shown to be due to a mutation in *Hoxa13*, the human homologue *HOXA13* immediately became a candidate gene for HFGS, and indeed mutations were identified (Mortlock and Innis, 1997).

However, mouse models of disease sometimes show considerably different phenotypes from the corresponding human disorders. Some difference in biochemical and developmental pathways in mice and humans are known and often transgenic or knockout (KO) mice show a dissimilar phenotype from the human disease. Phenotypic differences may be due to an altered pattern of transgene expression, an altered or

absent developmental pathway, a redundant or altered biochemical pathway or an unknown mechanism of mutation (Wynshaw-Boris, 1996). Modifier genes in different strains of mice could sometimes be responsible for differences in phenotype between mouse and human mutants.

A significant number of spontaneous mutations in mice cause congenital limb malformations and have been proven to be crucial to our understanding of the molecular mechanisms regulating limb bud morphogenesis.

1.2.3 The *limb deformity* mouse

The limb abnormalities in CLS patients were noted to be very similar to those found in a recessive mouse mutant, *limb deformity* (*ld*) (Winter, 1988). Homozygous *ld/ld* mice have fully penetrant, symmetrical distal limb abnormalities, including soft tissue and bony syndactyly, oligodactyly, fusion of carpals and metacarpals in the paws and of tarsals and metatarsals in the feet together with synostosis of the long bones of the zeugopod (radius and ulna, tibia and fibula) (Woychik *et al.*, 1985). In CLS, the upper limbs are more severely affected than the lower limbs and tibia-fibula synostosis is not usually a feature. In addition *ld/ld* mice often show partially penetrant and asymmetrical renal agenesis or hypoplasia (Maas *et al.*, 1994), which are not typical characteristics of CLS.

1.2.3.1 Background on Formin-1 role in ld phenotype

At the commencement of this investigation, and until very recently, the gene underlying the *ld* mouse phenotype was thought to be *Formin-1*. Identified in 1985, after chance insertion of a transgene resulted in a mouse with *ld* phenotype (Woychik *et al.*, 1985), the gene was originally called *Formin*, but after a closely related mouse gene, *Formin-2*, was discovered, *Formin* has been re-named *Formin-1* (Leader and Leder, 2000). Over the years, a total of five *ld* alleles have been identified. Two *ld* alleles (ld^{TgHd},

ld^{TgBri}) arose by chance insertional mutagenesis disrupting the C-terminal region of the *Formin-1* gene (Wang *et al.*, 1997). A third *ld* allele, ld^{In2}, was due to a ~40 Mb inversion with breakpoints in the C-terminal region of *Formin-1* (Woychik *et al.*, 1990a). The molecular lesions in the remaining two alleles, ld^{OR} and ld^{J}, have been identified only very recently and are discussed in more detail in section 1.2.3.2 (Zuniga *et al.*, 2004).

It had been assumed for a long time that the *ld* phenotype was caused by disruption of the C-terminal Formin-1 domain. In *ld* homozygous embryos, the epithelial-mesenchymal signaling interactions regulating limb bud development are disrupted. *Ld/ld* mice also have renal hypoplasia/agenesis, or hydronephrosis, resulting from failure or delay in the outgrowth of the ureteric bud. The model that was originally proposed, suggested that *Formin-1* played an important role in the reciprocal inductive interactions between mesenchyme and ectoderm in developing limb and kidney. Abnormalities in the kidney could be due to failure of inductive interactions between the ureteric bud and metanephric mesenchyme.

In the limb, *Formin-1* was thought to be essential for maintaining the feedback loop between the two signalling centres that co-ordinate outgrowth and patterning of the limb bud; namely, the apical ectodermal ridge (AER) and the zone of polarizing activity (ZPA) (section 1.3 and Fig. 1.3) (Zuniga and Zeller, 1999). The AER, a ridge of thickened ectoderm extending along the distal tip of the limb bud, promotes proximo-distal outgrowth of the limb by maintaining the mesenchyme immediately underneath it (called the progress zone) in a state of proliferation (Tickle, 2003). The ZPA, a specialised group of mesenchymal cells at the posterior margin of the limb, is important for controlling polarisation across the antero-posterior axis of the limb and establishing digit identity. Maintenance of both the AER and ZPA depends on a feedback loop between them. The signal from the ZPA is Sonic hedgehog (Shh); the signal from the AER comprises members of the fibroblast growth factor (Fgf) family. Genes encoding several different Fgfs (Fgf4, Fgf8, Fgf9, Fgf10 and Fgf17) are expressed in complex dynamic patterns during limb development and are now thought to act together to help maintain *Shh* expression (Moon *et al.*, 2000; Sun *et al.*, 2000; Boulet *et al.*, 2004).

In *ld/ld* mice, *Shh* expression in the ZPA is initially normal, but is not maintained and propagated, due to failure of *Fgf* expression in the AER (Zuniga *et al.*, 1999); this in turn is due to a mesenchymal defect in transducing the Shh signal to the AER, so that

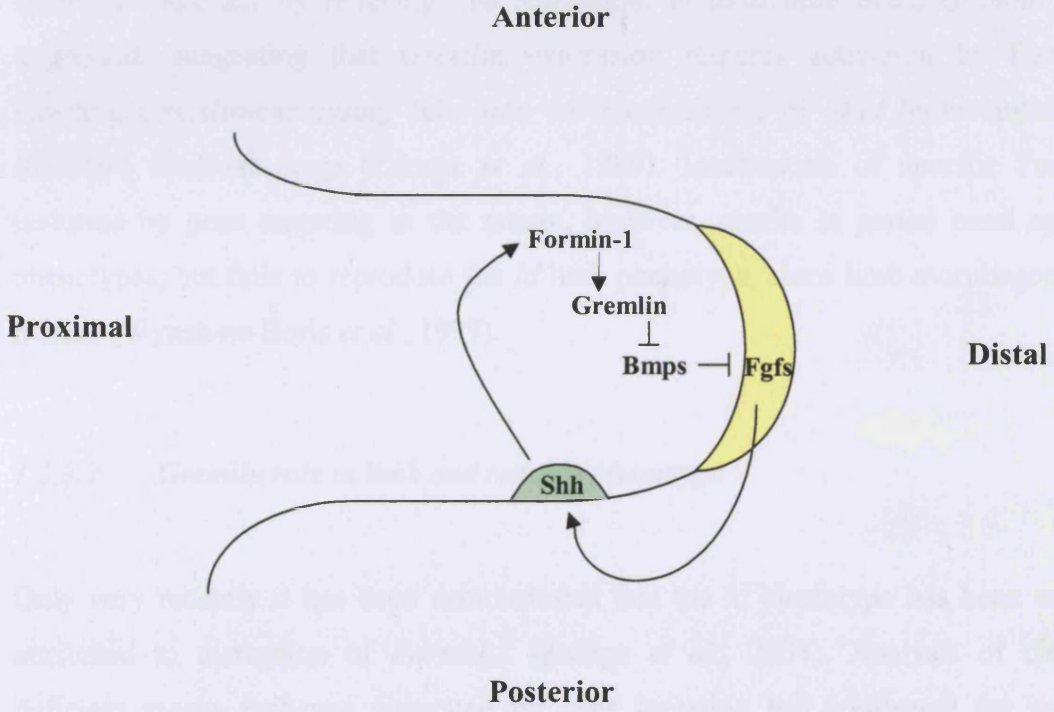


Figure 1.3 Schematic representation of the Shh-Fgf signaling loop and antagonistic interactions regulating limb development. The apical ectodermal ridge (AER) is indicated in yellow and the zone of polarising activity (ZPA) in green.

the AER fails to develop properly and is poorly organised and incompletely differentiated (Zeller *et al.*, 1989). Subsequently, it was suggested that mesenchymal Formin-1 was required to relay the Shh signal from the ZPA to the AER. A second mesenchymal signal downstream of Formin-1 then relays the signal on to the AER. This is Gremlin, a secreted signaling molecule which is a member of the DAN/Cerberus family and belongs to the TGFβ superfamily (Hsu *et al.*, 1998). Gremlin is a bone morphogenetic protein (BMP) antagonist that antagonises preferentially Bmp2 and Bmp4. Several BMPs are expressed in the mesenchyme and AER, and are thought to repress *Fgf* expression in the AER (Zuniga *et al.*, 1999). *Gremlin* could act by relieving this repression. In *ld/ld* limb buds, *Gremlin* is not expressed, suggesting that *Gremlin* expression requires activation by Formin-1. Grafting *Gremlin*-expressing cells into the mesenchyme of *ld/ld* limbs restores the *Shh/Fgf4* feedback loop (Zuniga *et al.*, 1999). Inactivation of specific Formin-1 isoforms by gene targeting in the mouse, however, results in partial renal agenesis phenotypes, but fails to reproduce the *ld* limb phenotype, since limb morphogenesis is normal (Wynshaw-Boris *et al.*, 1997).

1.2.3.2 Gremlin role in limb and renal ld phenotype

Only very recently it has been demonstrated that the *ld* phenotype has been wrongly attributed to disruption of *Formin-1* (Zuniga *et al.*, 2004). Analysis of *Gremlin*-deficient mouse embryos generated by gene targeting has confirmed the essential functions of *Gremlin* during limb bud development (Khokha *et al.*, 2003; Michos *et al.*, 2004). The Shh-Fgf feedback loop is disrupted in *Gremlin* mutant mice indicating an important role of Gremlin in maintaining the AER and the Shh-Fgf feedback loop during early limb development and patterning (Khokha *et al.*, 2003). Interestingly, it was noted that homozygous *Gremlin* KO mice have an identical phenotype to homozygous *ld* mice which results from defects in the *Formin-1* gene. The limb and kidney abnormalities were exactly the same as in *ld* mice. Furthermore, mice heterozygous for both the *Gremlin* null mutation and the ld^J mutation have the full *ld*

phenotype showing that the *Gremlin* null mutation is actually allelic to ld^{J}, one of the five previously identified *ld* alleles (Khokha *et al.*, 2003).

Michos *et al.* (2004) showed that in *Gremlin*-deficient embryos a distinct and functional AER fails to form. The disruption of AER function blocks propagation of *Shh* expression by the polarizing region as was observed in *ld* mutant limb buds (Michos *et al.*, 2004). In addition, a more general role of *Gremlin* in epithelial-mesenchymal signaling during organogenesis was described (Michos *et al.*, 2004). *Gremlin*-deficient embryos have complete renal agenesis revealing that Gremlin is required for the induction of metanephric kidney organogenesis. Furthermore, incomplete differentiation of lung airway epithelia in the mouse mutant causes neonatal lethality (Michos *et al.*, 2004).

The relationship between *Gremlin* and *Formin-1* is interesting because in mice, as in humans, they lie very close together, only 40 kb apart. Zuniga and colleagues identified the remaining two *ld* alleles, ld^{J} and ld^{OR}, and established that they are loss-of-function mutations directly disrupting the *Gremlin* gene product (Zuniga *et al.*, 2004). The ld^{J} mutation is a point mutation which affects splicing of the *Gremlin* transcript and thus results in the truncation of the 5' part of *Gremlin* exon 2, whereas the complete *Gremlin* ORF encoded by exon 2 is deleted by the ld^{OR} mutation. Furthermore, a shared *cis*-regulatory element within *Formin-1* genomic regions that is required for both *Formin-1* and *Gremlin* expression as well as for *Gremlin* activation in the posterior limb bud mesenchyme was also identified in this study (Zuniga *et al.*, 2004). This region was demonstrated to be the region either deleted or disrupted in the hypomorphic *ld* alleles ld^{TgHd}, ld^{In2} *and* ld^{TgBri}.

In conclusion, taken together these latest studies establish that the *ld* phenotype is a direct consequence of loss of *Gremlin* expression in the limb bud mesenchyme and not due to disruption of *Formin-1* functions as previously assumed (Zuniga *et al.*, 2004). It also shows that despite early speculations (section 1.2.3.1) *Gremlin* and *Formin-1* are not part of the same pathway, although these two genes are coexpressed during organogenesis in various embryonic tissues such as kidney (Michos *et al.*, 2004; Wynshaw-Boris *et al.*, 1997).

1.2.4 Formin-1

1.2.4.1 Formin-1 gene

Formin-1 is a very large and complex gene that spans over 400 kb and contains at least 24 exons on mouse chromosome 2 (Wang *et al.*, 1997). Some introns are as large as 50 kb. *Formin-1* has four major splice forms expressed in a variety of fetal and adult tissues (brain and testis, as well as limb and kidney) (Zeller *et al.*, 1999). These isoforms share the same C-terminal region, but have isoform-specific N-termini. The promoter regions of the gene have not yet been characterised. There may be a promoter for isoform IV, lying within the gene upstream of exon 6 (Chan *et al.*, 1995; Jackson-Grusby *et al.*, 1992); but a promoter for isoforms I to III has not been identified in the 650 bp upstream of exon 1 (Jackson-Grusby *et al.*, 1992; Woychik *et al.*, 1990b). The limb bud *cis*-regulatory element shared by *Formin-1* and *Gremlin* lies within the genomic region encompassing *Formin-1* exons 19-23 (Zuniga *et al.*, 2004).

1.2.4.2 Formin-1 protein

Formin-1 is the founding member of a multigene family that mediates cytoskeletal rearrangements in response to signals that induce, for example, cell polarisation (section 1.2.4.3) (Evangelista *et al.*, 2003). The role of Formin-1 at a cellular/molecular level is not well understood. The subcellular site of Formin-1 activity is not yet clear. Early studies found that it localises to the nucleus (Trumpp *et al.*, 1992); a later study found that Formin-1 is both cytoplasmic and nuclear and that *ld* mutations cause it to be retained in the cytoplasm (Chan and Leder, 1996). It has been proposed that nuclear Formin-1 acts as a transcriptional regulator (Zeller *et al.*, 1999).

All Formin-1 isoforms have a proline-rich region, encoded by exon 9, called the FH1 domain. This domain matches consensus sequences for Src homology 3 (SH3) ligands.

One study used this proline-rich region to screen a mouse limb bud expression library for proteins that might interact with Formin-1 and found two different classes: proteins containing one or more SH3 domains and proteins containing a WW motif (Chan *et al.*, 1996). The proteins in this second class, which were all novel, were given the name 'formin binding proteins' (FBPs). However, there is no evidence so far that Formin-1 actually interacts *in vivo* with any particular protein containing a SH3 or WW domain (Leader and Leder, 2000).

Recently, Kobielak and colleagues, in an attempt to analyse alpha-catenin binding partners using a yeast 2-hybrid system, found direct interaction between alpha-catenin and Formin-1 (Kobielak *et al.*, 2004). They found that Formin-1 localises to adherens junctions and nucleates unbranched actin filaments *in vitro* and functions *in vivo* in alpha-catenin-dependent, radial actin cable formation. Disruption of the alpha-catenin-Formin-1 interaction blocks assembly of radial actin cables and perturbs intercellular adhesion.

1.2.4.3 Formin homology proteins

An increasingly large number of proteins closely related to Formin-1 have been discovered (Tanaka, 2000; Wasserman, 1998; Katoh, 2004a; Katoh, 2004b; Westendorf and Koka, 2004). They are called the formin homology (FH) proteins or formins. Over twenty are now known, isolated from a wide variety of species, including plants and fungi, as well as flies, worms and mammals. Their main role is in organising the cytoskeleton as actin nucleators in the formation of new filaments and they exert their effects on microtubule networks during meiosis, mitosis, the maintenance of cell polarity, cell migration, vesicular trafficking, signaling to the nucleus and embryonic development (Koka *et al.*, 2003; Wallar and Alberts, 2003). FH proteins are best defined by the presence of the unique, highly conserved, approximately 400 residue region, called the formin homology domain 2 (FH2) which has recently been found to nucleate actin filaments *in vitro* (Pring *et al.*, 2003; Zigmond, 2004). The FH1 domain

has been found to bind profilin, an actin monomer-binding protein which controls actin polymerisation and enhances nucleation.

Studies of FH proteins in yeast, have shown that these proteins are specifically required for the assembly of bundles of parallel actin filaments known as cables (Evangelista *et al.*, 2002; Sagot *et al.*, 2002; Evangelista *et al.*, 2003). All the defects associated with loss-of-function mutations in FH proteins, including defects in spindle formation, cytokinesis and cell polarity, could be due to failure of proper assembly of actin cables, which would then affect other cytoskeletal structures (Lew, 2002).

1.2.5 Gremlin

Gremlin, also known as CKTSF1B1 (cysteine knot superfamily 1, BMP antagonist 1), is a member of the DAN/Cerberus family of BMP antagonists, highly conserved through evolution and known to bind and block BMP2, BMP4 and BMP7 *in vivo* and *in vitro* (Hsu *et al.*, 1998). Gremlin contains a highly conserved cysteine rich repeat region called a cysteine knot. BMPs constitute a large family of secreted growth factors belonging to the TGFβ superfamily which are known to be expressed during embryonic limb development. Their functions include control of mesodermal cell proliferation, chondrogenic differentiation, induction of apoptosis, regulation of growth and regression of the AER and regulation of the anteroposterior axis of the early limb bud by exerting a negative effect on the Shh/Fgf feedback loop (Merino *et al.*, 1999; Zuniga *et al.*, 1999). A number of BMP antagonists such as Noggin, DAN, Chordin and Follistatin, spatially and temporally regulate different specific BMPs during limb patterning, preventing their interaction with their receptors. In particular, Gremlin expression is restricted to the distal limb bud mesenchyme and concentrated posteriorly. Gremlin has a major role in maintaining the AER and the Shh/Fgf feedback loop, as outlined in sections 1.2.3.1 and 1.2.3.2 (Zuniga *et al.*, 1999; Zuniga *et al.*, 2004). A potential role of Gremlin in processes such as glomerulosclerosis, tubulointerstitial fibrosis and cellular hypertrophy has been described (Dolan *et al.*,

2003). More recently, *Gremlin* expression has been demonstrated in a variety of embryonic structures including lung and kidney rudiments (Michos *et al.*, 2004).

1.3 General aspects of limb development

A limb consists of four segments: a root (zonoskeleton), a proximal segment (stylopodium) consisting of a single bone (humerus, femur), a medial segment (zeugopodium) consisting of two bones (radius and ulna, tibia and fibula), and a distal more complex part (autopodium - hand and foot) (Fig. 1.4A). Correct formation of these segments requires coordination of several processes (Tickle, 2004). The limb develops from the embryonic limb bud through a rapid cell proliferation process and its development is mediated by reciprocal interaction between specific signaling regions within the limb bud. A long-standing model for limb development has been the progress zone model. The zone of undifferentiated mesenchyme that is preserved at the limb bud tip during outgrowth was named the progress zone and it was thought that the amount of time mesenchyme cells spend in this zone determined whether they form proximal or distal structures (Tickle, 2003). Recently this model has been challenged and an alternative model has been proposed according to which all the structures along the proximo-distal axis are already specified in the early limb bud and not progressively in a sequence linked to cell differentiation (Dudley *et al.*, 2002; Sun *et al.*, 2002).

1.3.1 Molecular basis of limb development

In both progress zone and early specification models, limb bud outgrowth is essential for establishing the pattern along the proximo-distal axis. A number of reciprocal ephithelial-mesenchymal interactions between the thickened apical ectodermal ridge (AER) at the distal tip of the limb bud and the underlying mesenchyme mediate limb bud outgrowth (Fig. 1.4B). Dorso-ventral patterning of the limb is controlled by signaling of the ectoderm covering the side of the bud (Fig. 1.4C), whereas the position

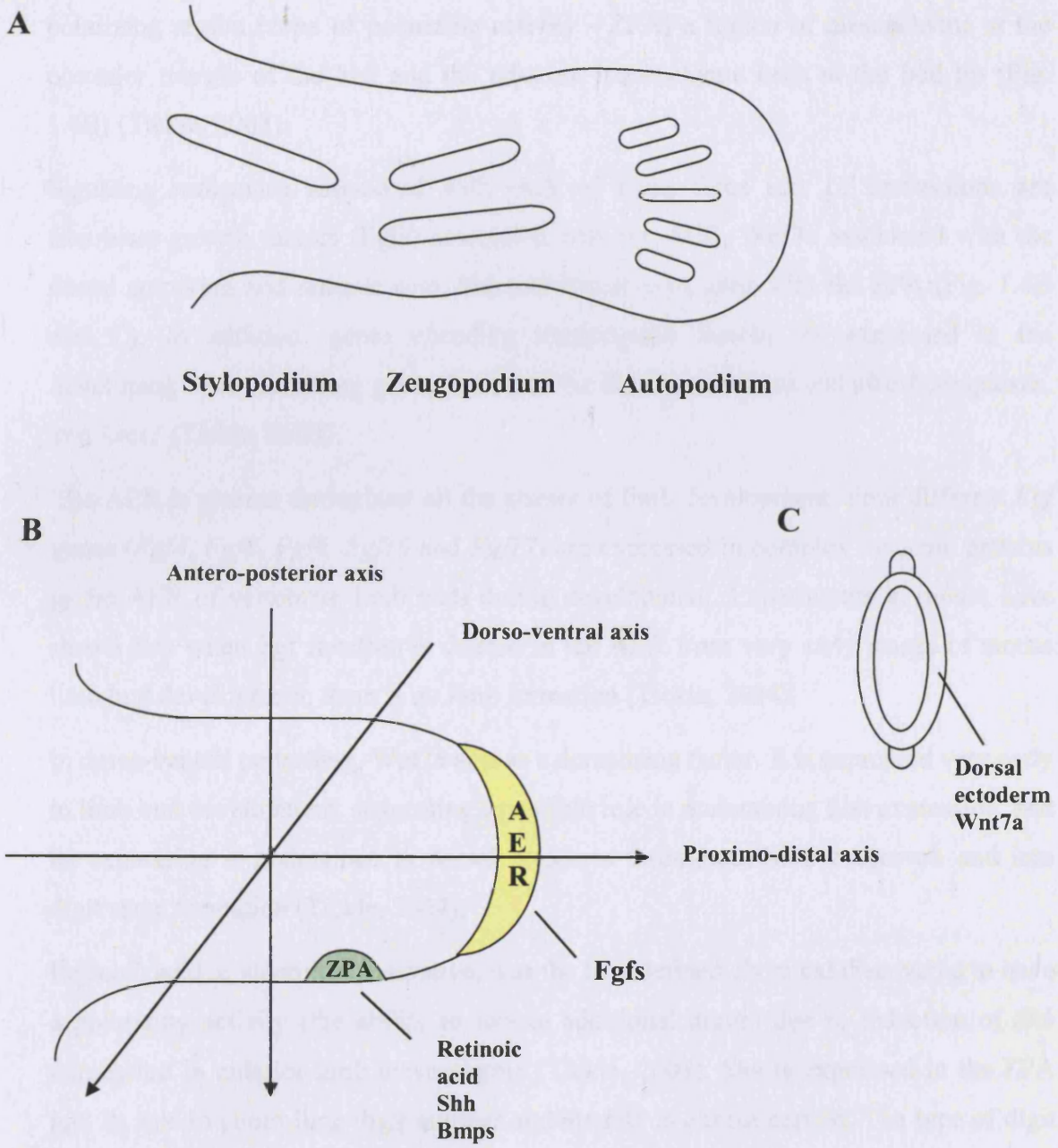


Figure 1.4 Schematic representation of the main segments of the limb (A) and view of limb patterning in relation to the three main axes of the limb (B). The apical ectodermal ridge (AER), the zone of polarising activity (ZPA) and the associated signalling molecules are indicated. A cross-section view of the limb bud and the dorsal ectoderm with the associated signal is also shown (C).

along the antero-posterior axis is specified by a signaling interaction between the polarizing region (zone of polarizing activity - ZPA) a region of mesenchyme at the posterior margin of the bud and the adjacent mesenchyme cells at the bud tip (Fig. 1.4B) (Tickle, 2003).

Signaling molecules associated with each of these three sets of interactions are fibroblast growth factors (Fgfs) associated with the AER, Wnt7a associated with the dorsal ectoderm and retinoic acid, Shh and Bmp2 associated with the ZPA (Fig. 1.4B and C). In addition, genes encoding transcription factors are expressed in the developing limb, including genes from the *Tbx* family, the *Hoxa* and *Hoxd* complexes, and *Lmx1* (Tickle, 2003).

The AER is present throughout all the phases of limb development. Four different *Fgf* genes (*Fgf4*, *Fgf8*, *Fgf9*, *Fgf10* and *Fgf17*) are expressed in complex dynamic patterns in the AER of vertebrate limb buds during development. Experiments in mouse have shown that when Fgf function is deleted in the AER from very early stages of mouse limb bud development, there is no limb formation (Tickle, 2004).

In dorso-ventral patterning, Wnt7a acts as a dorsalising factor. It is expressed very early in limb bud development, suggesting a possible role in maintaining Shh expression, and its expression is maintained in dorsal ectoderm throughout limb outgrowth and into digit stage formation (Tickle, 2004).

Retinoic acid, a vitamin A derivative, was the first defined chemical discovered to have a polarizing activity (the ability to induce additional digits) due to induction of *Shh* expression in anterior limb mesenchyme (Tickle, 2003). *Shh* is expressed in the ZPA and its role in controlling digit number and identity is almost certain. The type of digit that develops in a particular position, digit identity, is related to the strength of Shh signaling with the highest level of signaling being required for specification of the most posterior digit. In absence of Shh signaling, structures distal to the elbow/knee are very reduced.

Shh signaling in the limb bud is mediated by the zinc finger transcription factor Gli3. In the limb bud there is a gradient of repressor and activator forms of Gli3. Shh prevents processing of Gli3 to its repressor form and diffusion of Shh from the polarizing region

creates a concentration gradient across the bud. This results in a dose-dependent relief of inhibition imposed by high levels of the repressor form of Gli3 and thus allows expression of target genes with the downstream consequences of specification of digit number and specification of antero-posterior position leading to digit identity (Tickle, 2003). Shh signal specifies digit identity acting via bone morphogenic proteins (Bmps), in particular Bmp2, which acts downstream of Shh. *Bmp2* is expressed posteriorly in normal limb buds but whether it represents the polarizing morphogen or not is still controversial. Another main downstream target of Shh signaling is an apical ridge maintenance factor produced by the ZPA in order to specify digit number. This is the Bmp antagonist Gremlin, whose expression is inhibited by high levels of the repressor form of Gli3. Gremlin antagonises Bmp signaling in the AER thus allowing expression of Fgf4 and the AER function to be maintained (section 1.2.3.1)(Zuniga *et al.*, 1999).

Members of the *Tbx* family of genes encoding transcription factors have recently been shown to be implicated in digit and limb identity. In chick embryos, posterior expression of *Tbx2* and *Tbx3* regulates posterior digit identity, whereas *Tbx4* and *Tbx5* are expressed in regions of the embryos that will give rise to legs and wings respectively (Tickle, 2004).

Finally, different combinations of *Hoxa* and *Hoxd* genes are known to be responsible for patterning the different limb segments. *Hoxd13* and *Hoxa13* pattern the digits, *Hoxd11* and *Hoxa11* the lower arm/leg, *Hoxd9* and *Hoxa9* the upper arm and *Hox10* paralogs the upper leg (Tickle, 2004). *Hox* gene expression changes in the developing limb are induced by a combination of Fgf signaling with retinoic acid, Shh, and Bmps. The 5' *Hoxd* genes are expressed very early in lateral plate mesoderm, while the 5' *Hoxa* genes are activated in sequence initially in lateral plate mesoderm and later into the limb bud. 5' genes in both clusters are also expressed at a later stage during patterning of digits.

Determination of the number of digits and their identity is mediated by cell-cell interactions in the developing limb bud. Both Bmp and Fgf signaling in mesenchyme and AER are known to be implicated in these processes. Separation of individual digits is then mediated by cell death of the soft tissue web between the digits, which is regulated by Bmps (Tickle, 2004).

1.3.2 Limb anomalies in humans

In humans, limb malformations occur in approximately one in a thousand neonates. Two classification schemes are currently favored by clinicians, the Temtamy and McKusack classification purely based on clinical descriptions (Temtamy and McKusick, 1978), and the Winter and Tickle classification based on embryological and clinical concepts that distinguish abnormalities related to defects in embryological patterning from those that are not (Winter and Tickle, 1993). Nowadays, the increasing knowledge of the genetic basis of many limb malformations enables classification to be made by their underlying genetic defects (reviewed in Wilkie, 2003). The list of genes that are known to be responsible for human limb defects continues to grow. Amongst the first genes to be discovered were the fibroblast growth factor receptors, mutations in which are responsible for achondroplasia and several other conditions with digit anomalies such as Apert syndrome (Wilkie, 2003). Mutations in the *GLI3* gene are responsible for Grieg cephalopolysyndatctyly (Wild *et al.*, 1997). Mutations in *HOXD13* were found to be associated with human synpolydactyly (Muragaki *et al.*, 1996), whereas defects in *HOXA13* have been identified in hand-foot-genital syndrome (Mortlock and Innis, 1997). *TBX* genes have also been implicated in Holt-Oram syndrome (*TBX5*) (Basson *et al.*, 1997; Li *et al.*, 1997) and ulnar-mammary syndrome (*TBX3*)(Bamshad *et al.*, 1997). More recently, mutations in regulatory elements of *SHH* have been associated with preaxial polydactyly (PPD)(Lettice *et al.*, 2003).

1.4 General aspects of kidney development

The kidney develops from the intermediate mesoderm, a strip of cells along the posterior wall of the abdominal cavity which connects the somite (dorsal mesoderm) to the lateral plate mesoderm (Fig. 1.5A). The mesoderm which will form the kidney is

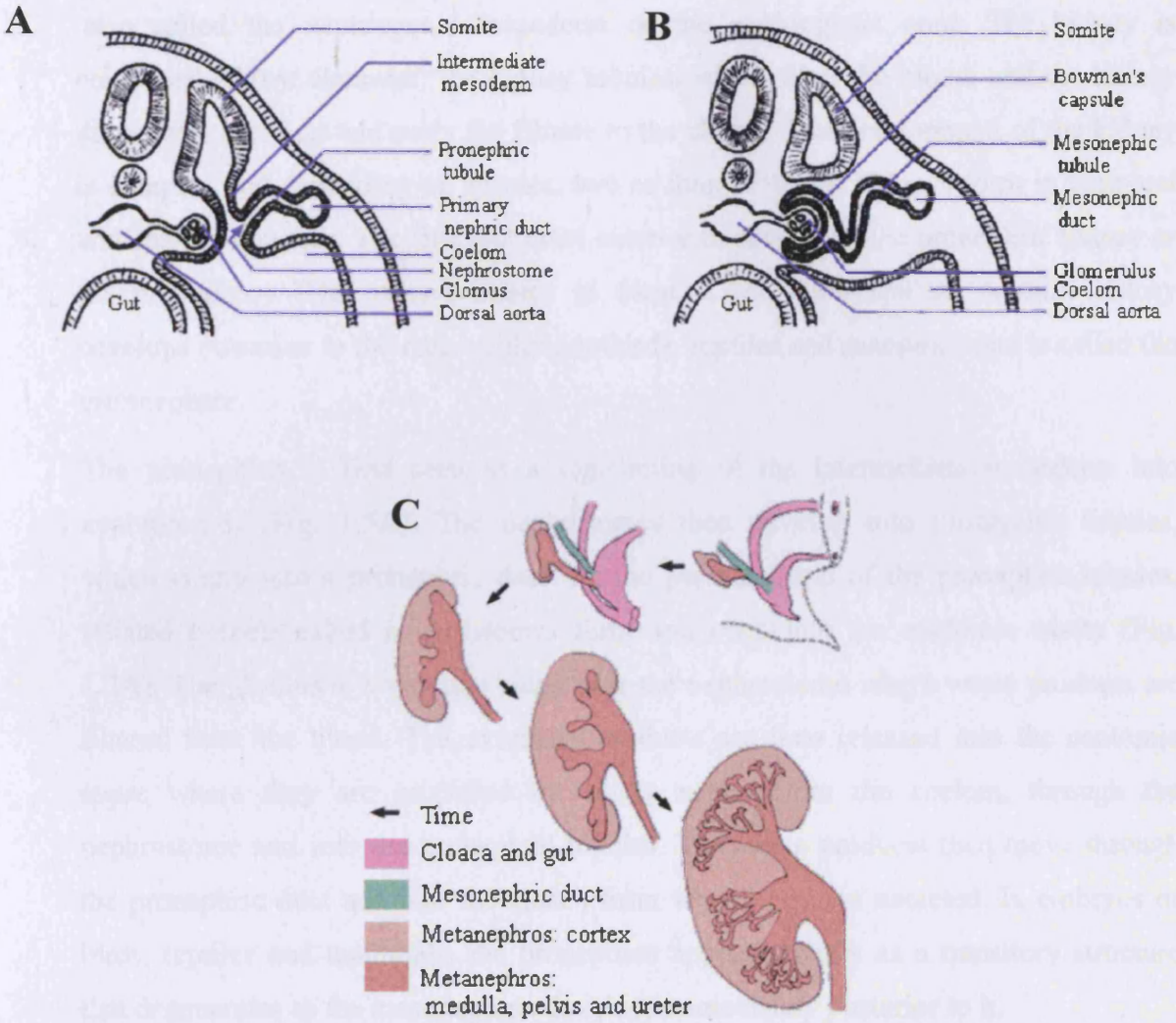


Figure 1.5 Schematic of kidney development. A transverse section showing the pronephric (A) and mesonephric (B) tubules is represented. The development of the metanephros to adult kidney is also shown (C). *Adapted with permission from http://www.uoguelph.ca/zoology/devobio/210labs/kidney2.html*

also called the nephrogenic mesoderm or the nephrogenic cord. The kidney is composed of two elements: the kidney tubules, which filter the blood, and the kidney ducts, which collect and carry the filtrate to the cloaca. The development of the kidney is complex and depending on species, two or three different kidneys form in temporal and spatial sequence. The first and most anterior to develop is the pronephric kidney or the pronephros. The second kidney to form is the mesonephros. A third kidney develops posterior to the mesonephros in birds, reptiles and mammals and is called the metanephros.

The pronephros is first seen as a segmenting of the intermediate mesoderm into nephrotomes (Fig. 1.5A). The nephrotomes then develop into pronephric tubules, which empty into a pronephric duct. At the proximal end of the pronephric tubules, ciliated funnels called nephrostomes form and open into the coelomic cavity (Fig. 1.5A). The glomus is a vascular ridge near the nephrostome where waste products are filtered from the blood. The excretory products are then released into the coelomic space where they are propelled by ciliary action from the coelom, through the nephrostome and into the pronephric tubules. The waste products then move through the pronephric duct and into the cloaca from where they are excreted. In embryos of birds, reptiles and mammals, the pronephros appears briefly as a transitory structure that degenerates as the mesonephros develops immediately posterior to it.

When the second embryonic kidney, the mesonephros, develops, the mesonephric tubules form an expanded, blind-ended region that invaginates to form a double walled cap called the Bowman's capsule (Fig. 1.5B). This invagination is caused by a group of capillaries, called a glomerulus, pushing into and becoming surrounded by the Bowman's capsule (Fig. 1.5B). The structure formed by the glomerulus surrounded by the Bowman's capsule is called a renal corpuscle or Malpighian corpuscle. Meanwhile, the mesonephric tubules become long and coiled with their distal ends opening into the mesonephric duct (also called archinephric ducts or Wolffian ducts).

The third kidney, the metanephros, develops from a diverticulum, called the ureteric bud, of the mesonephritic duct posterior to the mesonephros (Fig. 1.5C). The ureteric buds elongate into the region of presumptive metanephrogenic mesenchyme and induce the mesenchyme to form the nephric renal tubules. Reciprocal interactions between the

mesenchyme and the ureteric bud cause the latter to branch and form a tree-like system of collecting ducts (Fig. 1.5C). The initial tube portion of the ureteric bud is called the ureter. At the metanephric end of the ureter, an expansion called the renal pelvis forms and branches numerous times to form the collecting ducts. One end of the nephric tubules will connect to the collecting ducts, while the other end will form the Bowman's capsule that surrounds the glomerulus. The nephritic tubules become long, S-shaped and develop specialized regions. Subsequently, the ureters separate from the mesonephric ducts, elongate and serve to carry urine from the kidney to the bladder.

The adult kidney is bean shaped with the indent being called the hilum where the renal artery and nerves enter and the ureter and renal vein exit (Fig. 1.6A). The renal pelvis is immediately adjacent to the hilum. The periphery of the kidney is surrounded by a connective tissue capsule and beneath the capsule is a region called the cortex (Fig. 1.6B). The cortex is where the renal corpuscles (glomeruli plus Bowman's capsules) and most of the nephric tubule is located. Below the cortex is the medulla composed of nephric tubules, collecting ducts and numerous capillaries (Fig. 1.6B). The medulla narrows into a pyramid-shaped structure, the renal papilla, near the renal pelvis. Urine is discharged from the collecting ducts, which open into the renal papilla, and moves through the renal papilla to the renal pelvis and out of the kidney through the ureter.

1.4.1 Molecular basis of kidney development

Although the morphogenetic steps of kidney development have been well characterised, the nature of many of the molecules involved in this process are still unknown. However, the use of genetically modified mice, zebrafish and *Xenopus laevis* has provided valuable models to elucidate the role of some of these molecules. There are a growing number of potential molecules that are expressed during embryonic kidney development including growth factors and their receptors, transcription factors, integrins, extracellular matrix components and proteoglycans.

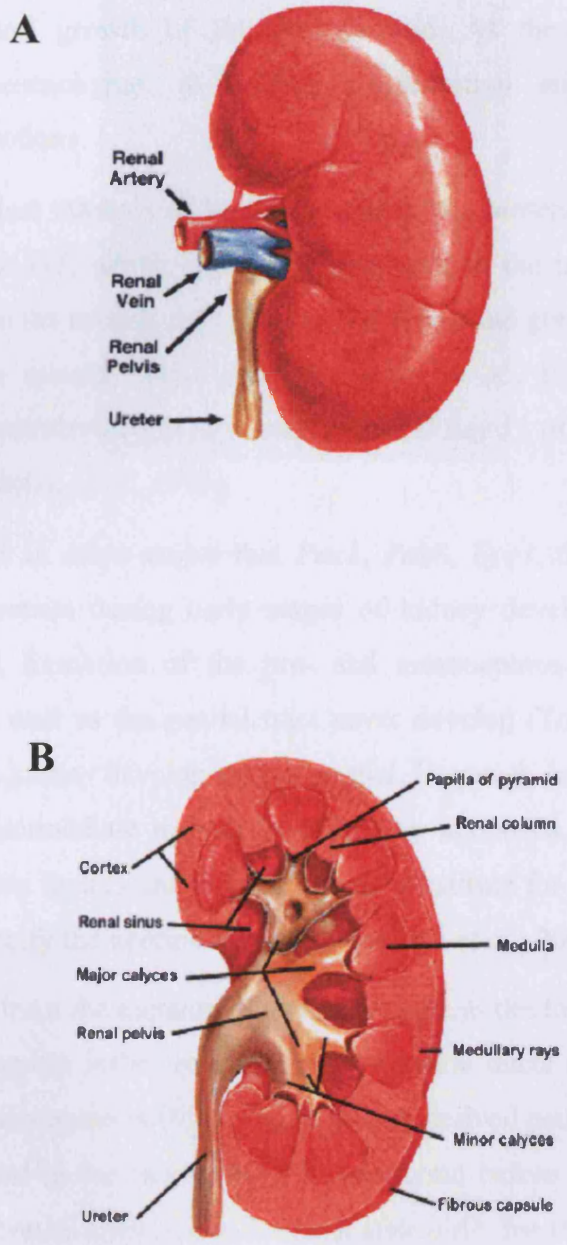


Figure 1.6 Gross morphology of the kidney (A and B). *Adapted with permission from http://www.auburn.edu/~deruija/renal_part1 /sld006.htm*

It has been proposed that development of the definitive metanephric kidney depends on a series of sequential and reciprocal molecular interactions between the ureteric bud epithelium and the metanephric mesenchyme. The development of the kidney initiates by formation and growth of the ureteric bud. As the ureteric bud invades the metanephric mesenchyme, it induces condensation and nephrogenesis through reciprocal interactions.

One of the earliest markers of the metanephrogenic mesenchyme is the Wilms tumor suppressor gene *Wt1*, which is initially expressed in the intermediate mesoderm and then restricted to the metanephrogenic mesenchyme and gonads, amongst other tissues, but absent from ureteric bud tissue (Armstrong *et al.*, 1993). In *Wt1* null mice the metanephric mesenchyme can be identified, but at day 11 of gestation the cells undergo apoptosis (Kreidberg *et al.*, 1993).

Genetic analysis in mice shows that *Pax2*, *Pax8*, *Eya1*, *Six1* and *Six2* transcription factors are expressed during early stages of kidney development. In *Pax2*-deficient embryos, initial formation of the pro- and mesonephros occurs normally, but the metanephros as well as the genital tract never develop (Torres *et al.*, 1995). In *Pax8* mutant embryos kidney development is normal. However, in Pax2, Pax8 double-mutant embryos, the intermediate mesoderm is lost by apoptosis, suggesting that Pax2 and Pax8 transcription factors can at least in part substitute for each other in development and together specify the nephric lineage (Bouchard *et al.*, 2002).

The first signal from the metanephric mesenchyme in the form of a diffusible signaling molecule causing the initial budding of the nephric ducts and elongation of the buds towards the mesenchyme is Gdnf (glial-cell-line-derived neurotrophic factor). The *Gdnf* gene is expressed in the metanephric mesenchyme before ureteric bud induction and mice with inactivated *Gdnf* alleles die soon after birth due to renal agenesis (Sanchez *et al.*, 1996). Gdnf promotes the formation of primary ureteric buds acting as a ligand for the c-Ret receptor tyrosine kinase that is located at the tip of the ureteric bud (Trupp *et al.*, 1996). *Gdnf* expression is regulated by a variety of different factors, which have been recently identified through gene knockout experiments. Genetic ablation of the *Drosophila sine oculis* homolog *Six1*, for example, leads to a variety of kidney phenotypes including agenesis (Li *et al.*, 2003). The ureteric bud forms but does not

invade the metanephric mesenchyme completely. There is no induction of nephron tubules and the metanephric mesenchyme undergoes apoptosis (Li *et al.*, 2003). In *Six1* mutant metanephric mesenchyme, *Gdnf* expression is reduced. SIX1 protein binds the *Gdnf* promoter in a complex with EYA1 protein tyrosine phosphatase, suggesting a direct role for *Six1* in transcriptional regulation of *Gdnf* (Li *et al.*, 2003). *Eya1* knockout mice show loss of *Gdnf* expression as well as loss of *Six1* and *Six2* expression.

The *Sal-like 1* (*Sall1*) deficient mouse also shows defects in full invasion of the ureteric bud into the metanephric mesenchyme (Nishinakamura *et al.*, 2001). In *Six1* mutants, *Sall1* expression is lost, suggesting that *Six1* may act upstream of *Sall1*.

The transforming growth factor family member 11 (*Gdf11*) is expressed in both the Wolffian duct and the metanephric mesenchyme. *Gdf11* null mutant kidneys exhibit bilateral agenesis, unilateral agenesis or hypoplasia (Esquela and Lee, 2003). *Gdnf* expression is not detected in the mutant mesenchyme, suggesting that *Gdf11* may regulate *Gdnf* expression and therefore ureteric bud outgrowth. Alternatively, loss of *Gdnf* expression could result secondarily from failure of ureteric bud invasion.

A role in restricting the *Gdnf* expression domain is played by the secreted Slit2 protein and its Robo2 receptor. In both *Slit2* and *Robo2* mutants, multiple ureteric buds form anterior to where normal ureteric bud outgrowth occur, and multilobar kidneys are observed (Grieshammer *et al.*, 2004). The anterior boundary of the *Gdnf* expression domain in *Slit2* and *Robo2* mutants is expanded and reducing *Gdnf* expression rescues the mutant phenotype. *Eya1* and *Pax2* expression are unchanged in mutants, suggesting that *Slit2/Robo2* does not regulate *Gdnf* expression through these molecules.

Recently, a number of ureteric bud-expressed genes have been identified and described to be important in branching of the ureteric bud and rescuing the metanephric mesenchyme from apoptosis after the outgrowth of the ureteric bud. The *FRAS1* gene, for example, the gene responsible for Fraser syndrome (a congenital disorder affecting several systems including the kidney) has been recently mapped (McGregor *et al.*, 2003). Studies in the mouse mutant model for Fraser syndrome, *blebbed*, in which the homolog of *FRAS1* was found to be truncated by mutation, and in *Fras1* mutants

suggest that Fras1 may play a role in local interactions that regulate the survival and the induction of the metanephric mesenchyme (Vrontou *et al.*, 2003). In both *blebbed* and *Fras1* mutants, in fact, the ureteric bud forms and invades the metanephric mesenchyme, but induction does not occur resulting in apoptosis of the mesenchyme and termination of ureteric bud growth. In another mouse model of Fraser syndrome, *eyeblebs*, the *Grip1* gene is disrupted (Takamiya *et al.*, 2004). Gene targeting of *Grip1* showed complete renal agenesis. The ureteric bud forms and invades the metanephric mesenchyme, but induction of the mesenchyme fails to occur as shown by a lack of Wt1 upregulation and an increase in apoptosis. *Grip1* and *Fras1* colocalize to the basal surface of epithelial cells and they can physically interact.

As outlined in section 1.2.3.2, it has been recently demonstrated that the BMP antagonist *Gremlin* is required to initiate ureter growth and branching that in turn induces metanephric nephrogenesis (Michos *et al.*, 2004). *Gremlin* is initially expressed in the intermediate mesenchyme and subsequently in the posterior Wolffian duct and metanephric mesenchyme. During branching, *Gremlin* is expressed locally in the mesenchyme surrounding the invading ureter. In *Gremlin* mutants, the ureteric bud forms, but fails to invade the metanephric mesenchyme. In the mutant metanephric mesenchyme, Gdnf expression is reduced at embryonic day 11.25 and lost by embryonic day 11.5 resulting in failure of branching. Pax2 expression is also lost by embryonic day 12.5 suggesting a possible failure to induce condensation of the metanephric mesenchyme.

1.5 Strategies for gene identification

To identify a disease gene four main strategies are employed: functional cloning, positional cloning, the candidate gene approach or a combination of the latter, the positional candidate approach (Collins, 1995). Many varied pathways have been used to identify disease genes, but ultimately all methods aim to identify one or more plausible candidate genes, which are then examined further for evidence that they associate with the disease.

1.5.1 Functional cloning

A gene can be identified by a position-independent strategy with no prior knowledge of its chromosomal position (Ballabio, 1993; Collins, 1992). This approach requires information about the function of the unidentified gene and isolation of the protein product of the gene. The amino acid sequence can be determined if adequate amounts of protein are available. Partially degenerate oligonucleotides coding for the protein can be synthesized and used to screen a cDNA library. The cDNA clone(s) is then used to screen a genomic library to enable characterisation of the gene from a genomic clone. An alternative approach is to raise antibodies to the protein of interest and use these to screen an expression cDNA library.

The main limitation of this approach is that biological information about the function of most gene products is simply not available. This method has been successful only in a few cases, for example in enzyme deficiencies such as phenylketonuria, known to be due to deficiency of the enzyme phenylalanine hydroxylase or Haemophilia A, due to deficiency of Factor VIII (DiLella *et al.*, 1986; Gitschier *et al.*, 1984) .

1.5.2 The candidate gene approach

The pure candidate gene approach relies on some knowledge of the type of gene anticipated to be involved in a particular disorder, but no prior information regarding its chromosomal position. The major limitation of this approach is that predictions of the biochemical functions of an unknown gene are often inaccurate in view of the complexity of the disease. In addition, screening for mutations in multiple candidate genes is very time-consuming. However there are several ways in which 'educated guesses' for good candidates can be made.

A candidate gene for a human disorder may be proposed on the basis of similarity of the phenotype to the phenotype of another human disorder for which the gene is known. The expression pattern or function of a gene may support the hypothesis that it is a candidate for a particular disorder. Animal models, as already outlined in section 1.2.2, whether spontaneous or artificially generated, can also suggest candidate genes for human disorders, since if the animal model has a similar phenotype to the human condition, it might result from mutations in the animal orthologue of the human disease gene. Although this approach can prove extremely valuable, there are limitations to the use of animal models as the phenotype in humans and mice may differ considerably despite mutations in orthologous genes.

1.5.3 Positional cloning

In this method, a disease gene is cloned from an initial approximate knowledge of its subchromosomal localization, with no prior knowledge of its sequence or biochemical function to suggest a likely candidate (Collins, 1995). The initial localization of the disease gene could be a chromosomal location of 10 Mb or more, which may have been identified by linkage analysis (section 1.6) or the presence of a visible cytogenetic abnormality such as a translocation, a duplication or a deletion associated with the disease in some patients.

The first step is to refine the candidate region as precisely as possible. An initial genome-wide screen for linkage is normally carried out using markers spaced at relatively wide intervals (often 10-15 cM) and will therefore define a large candidate region. More closely spaced polymorphic markers need to be analysed once a candidate region has been identified. Disease haplotypes are identified in individual pedigrees by haplotype analysis and recombinants are identified in order to define proximal and distal flanking markers. This defines a candidate region, which cannot be narrowed further except by finding new recombinants in other affected families. Linkage disequilibrium (section 1.6.6) when present, can help to narrow the candidate region further.

The next step is to construct an ordered contig of clones, often YACs (yeast artificial chromosomes), across the critical region. Nowadays, with the progress of the Human Genome Mapping Project, an ordered contig is likely to be available for most regions of the genome. If not, it is possible to identify the YACs by screening publicly available libraries. As YACs contain large inserts (often up to 1 Mb) they can be often unstable and prone to rearrangements, so other vectors such as PACs (P1 artificial chromosomes) or BACs (bacterial artificial chromosomes) with smaller insert sizes (100-300 kb) may be used to form the contig. BAC clone data can be accessed in many different ways, either directly from sequencing centers or from publicly available databases. Once a physical map of the interval with overlapping DNA clones has been built, these can be used as a substrate to identified novel candidate genes.

Several methods are then used to identify expressed sequences within these genomic clones. These methods include cDNA selection, cDNA library screening using genomic clones from the candidate region as probes, "zoo blotting" to search for evolutionarily conserved sequenced, identification of CpG islands which are often situated close to genes, exon trapping to identify genomic sequences flanked by functional splice sites and sequence analysis. A transcript map of all expressed sequences within the region is then generated, and the most likely candidate genes can then be screened for evidence that they are associated with the disease.

An example of the use of the positional cloning approach is the identification of *FRAS1* gene for Fraser syndrome (McGregor *et al.*, 2003). Homozygosity mapping narrowed

the candidate region to a 1.5 cM interval on 4q21 and then a positional cloning approach was used to identify suitable candidate genes. The human cDNA sequence of *FRAS*1 in the region was determined from the cDNA sequence used in a targeted gene knockout causing an identical phenotype to the blebbed mouse, which had been previously postulated to be a model for Fraser syndrome (McGregor *et al.*, 2003; Vrontou *et al.*, 2003).

The positional cloning approach is laborious, although in recent years as the human genome sequence progresses, practical approaches to the characterisation of genomic loci have changed and become more efficient and rapid using web-based resources. The wide availability of transcript maps and the gradual increase in the number of annotated genes allowed the positional cloning to be superseded by the positional candidate approach which nowadays represents the most widely used method of gene identification (Collins, 1995).

1.5.4 The positional candidate approach

The majority of genes have been identified using the positional candidate approach which combines both the pure positional cloning and the positional independent candidate approaches, since the most robust candidate genes are those that map to the same chromosomal region as the disease. This approach has been facilitated by the increased information and resources made available by the Human Genome Project. Therefore, once a candidate region is identified (usually by linkage analysis), attractive candidate genes mapping to this region can be selected from public databases, which contain rapidly increasing numbers of mapped genes often identified only as cDNA or ESTs. However, a candidate region often contains a considerable number of genes and it is essential to select appropriate candidate genes and prioritise them for further investigation. This demands careful consideration of information on their spatial and temporal pattern of expression (from Northern blotting, RT-PCR or *in situ* hybridisation of mRNA in tissue sections), likely function or functional domains, and

homology to other genes implicated in similar human phenotypes, model organisms or relevant natural mutants.

The identification of the gene for recessive Robinow syndrome represents an example of how the positional candidate approach was successfully adopted. The condition was mapped to 9q22 by homozygosity mapping in consanguineous families (Afzal *et al.*, 2000a). This region overlapped with the locus for autosomal dominant brachydactyly B. The identification of *ROR2* as the gene mutated in brachydactyly B and the observation that homozygous knockout mice (*Ror2* -/-) show mesomelic dwarfing made this a candidate for Robinow syndrome. Afzal and colleagues subsequently showed that *ROR2* mutations are associated with recessive Robinow syndrome (Afzal and Jeffery, 2003; Afzal *et al.*, 2000b).

1.6 Linkage analysis

1.6.1 Recombination and linkage

The positional cloning and positional candidate approaches outlined above in sections 1.5.3 and 1.5.4 both rely on the identification of the chromosomal region of interest as a starting point. When there are no cytogenetic 'clues', such as a translocation, a duplication or a deletion, the method most commonly used to achieve the initial localisation is linkage analysis. The principle behind linkage analysis is that two loci that lie close to each other on the same chromosome (i.e. are syntenic) tend to segregate together during meiosis and thus tend to be inherited together as a 'linked unit', whereas the further apart two syntenic loci lie on a chromosome, the more likely it is that they will be separated by a recombination event during meiosis. If alleles at two loci cosegregate together in significantly more than 50% of all the observed meioses, then the two loci are said to show genetic linkage.

The recombination fraction (θ) is the measurement of genetic linkage. This decribes the chance of recombination occurring between two loci and is the probability that a parent

will produce a recombinant offspring. The recombination fraction (θ) is an indirect measure of the genetic distance between genetic loci and ranges from zero, for tightly linked loci, to 0.5 for loci that are unlinked or on different chromosomes (Terwilliger, 1994). The smaller the recombination fraction, the smaller the genetic distance between two loci. A recombination fraction of 0.01 (only one in every 100 offspring inherits the recombinant haplotype) corresponds to a genetic distance of one centimorgan (1 cM). A genetic distance of 1 cM approximates to a physical distance of 1 Mb, although this is variable as recombination is more likely to occur in some chromosomal regions than others, for example the telomeres, and is greater in females compared to males.

1.6.2 Lod scores

The lod score is defined as the logarithm to base 10 of the odds of linkage, and is a measurement of the likelihood of genetic linkage between two loci. It is a statistical score in which the likelihood of alleles at two genetic loci being inherited together because they are situated close together on the same chromosome (i.e. linked) is compared to the probability that this happened by chance (i.e. not linked). The lod score is therefore equal to the logarithm to base 10 of the ratio of the probability that two alleles are linked with a given recombination value θ divided by the probability that they are unlinked. The likelihood of observing a particular pattern of inheritance of a disease and a marker locus is calculated assuming varying degrees of linkage over a range of recombination fractions (θ = 0 to 0.5) and compared to the likelihood of observing the same inheritance pattern assuming no linkage (θ = 0.5).

The lod score is calculated according to the following formula:

$$\text{Lod score } (Z) = \log_{10} L(\theta = x) \;/\; L(\theta = 0.5)$$

where L is the likelihood of the observed inheritance pattern and x is a particular value of θ, usually between zero and 0.5. When the lod score reaches its maximum, the most likely recombination fraction is located. If there are no recombinants, the lod score will be maximal at θ = 0, but if there are recombinants, the lod score will be highest at the

most likely recombination fraction. This is a two-point lod score as it refers to linkage between two loci, i.e. the disease locus and the marker locus. There is usually enough evidence for linkage between two loci if the lod score is greater than 3 for a θ value. A lod value of 3, in fact, means that the chances are greater than 1000:1 that the two loci are linked than the odds that they are not. If the lod score is below –2, this is accepted as evidence excluding linkage to the region. Values in between –2 and +3 are inconclusive (Terwilliger, 1994).

A limitation in using linkage mapping in humans is that families are often too small to generate statistically significant linkage information. However, since lod scores are logarithms, it is possible to add together lod scores from separate small families. Human linkage analysis relies on computer programs such as LIPED or MLINK, which is part of the LINKAGE package (Terwilliger, 1994) to compute the lod scores.

Multipoint linkage analysis is an extension of the two-point linkage analysis. Instead of finding the recombination fraction between a single marker and a disease locus, multipoint analysis aims to map a disease locus against several markers at the same time and establishes between which two markers the gene is most likely to be located. The MLINK and LINKMAP programs from the LINKAGE package can both be used for multipoint analysis (Terwilliger, 1994).

1.6.3 Polymorphic markers

Multiple highly informative DNA polymorphic markers are used in a genome wide screen for genetic linkage, to see whether a marker at a particular locus segregates with the disease in a family. The earliest polymorphisms used in linkage analysis were enzyme polymorphisms and blood group markers.

Restriction fragment length polymorphisms (RFLPs) were the first type of DNA marker used in linkage analysis, but a whole genome scan using this method was extremely time consuming as it involved typing the RFLPs using radiolabelled probes hybridised to Southern blots of restriction digests. Moreover they had limited informativeness, as restriction enzyme sites are either present or absent and therefore

they give rise to two alleles. Minisatellite markers (or variable number tandem repeats, VNTRs) represented the next generation of markers. They had greater informativeness, but the technology available at the time still required the use of Southern blotting and hybridisation with radiolabelled probes.

The informativeness of a marker is measured by the polymorphism information content (PIC), which is a mathematical formula showing the probability of identifying the parental chromosome contribution in any given offspring. The PIC is calculated according to the following formula:

$PIC = 1- (p^2 + q^2 + r^2) - (2p^2q^2r^2)$ where p, q and r are allele frequencies.

It was only with the advent of PCR and the identification of multiple polymorphic microsatellite markers (mainly short $(CA)_n$ repeats) that mapping had been greatly facilitated allowing large-scale rapid genome-wide linkage analysis to be developed. Tri- and tetranucleotide repeats are increasingly being used for their advantage to be less prone to 'stutter' bands when amplified by PCR, which makes the allele size hard to interpret. Panels of fluorescent-labelled primers to amplify markers spaced at 10-20 cM density across the genome, which can be multiplexed, are available for high throughput genotyping. Detailed genetic linkage maps of microsatellites throughout the genome have been generated and are available on public databases (Dib *et al.*, 1996).

Nowadays with the advent of DNA microarray technology, single nucleotide polymorphisms (SNPs) represent a powerful high throughput tool for linkage analysis. Although SNPs are less informative than the currently used genetic markers, having only two alleles, they are much more abundant and widespread throughout the genome. More than 10,000 SNPs can be genotyped simultaneously per single DNA sample enabling a genome-wide screen to be performed rapidly. Efforts have been made to integrate information from existing genetic and physical maps to provide a high resolution map which includes SNP positions and is now publicly available (dbSNP at the NCBI website).

1.6.4 Homozygosity mapping

Recessive disorders are usually rare in the general population, but common in consanguineous families. Rare recessive diseases arise relatively more in consanguineous families usually because the affected children have inherited the same disease allele from each of their parents. Each of the parents has in turn received this disease allele from a recent common ancestor.

Homozygosity mapping is a method of mapping the genes underlying rare recessive disorders using consanguineous families. This strategy was first proposed by Lander and Botstein in 1987 (Lander and Botstein, 1987). The basic principle is that affected offspring of consanguineous parents are homozygous not just for the same disease allele, but they are also likely to be homozygous for a whole segment of the chromosomal region around the disease locus (Fig. 1.7). In each child of first-cousin parents, whether affected or not, on average 1/16 of the genome is expected to be identical-by-descent (IBD) (This is 1/64 for a child of second cousin parents; and 1/256 for a child of third cousin parents). Such regions can be detected by genotyping with polymorphic markers. If several closely-linked polymorphic markers are all homozygous, then the region is IBD. If just one isolated marker in a region is homozygous, and the flanking markers are heterozygous, then the isolated marker must just be identical-by-state (IBS), and the region as a whole is not IBD. If several different affected individuals from the same consanguineous family, and/or several affected individuals from different consanguineous families are analysed, it is very unlikely that they will all share the same region of IBD unless this region contains the disease gene. In homozygosity mapping, therefore, affected individuals from consanguineous families are genotyped with a set of polymorphic markers spanning the whole genome, in order to define a region where the affected individuals are all homozygous. This region is then highly likely to contain the disease locus.

This approach has several advantages over standard methods for carrying out linkage analysis in non-consanguineous families. Firstly, families consisting just of the parents

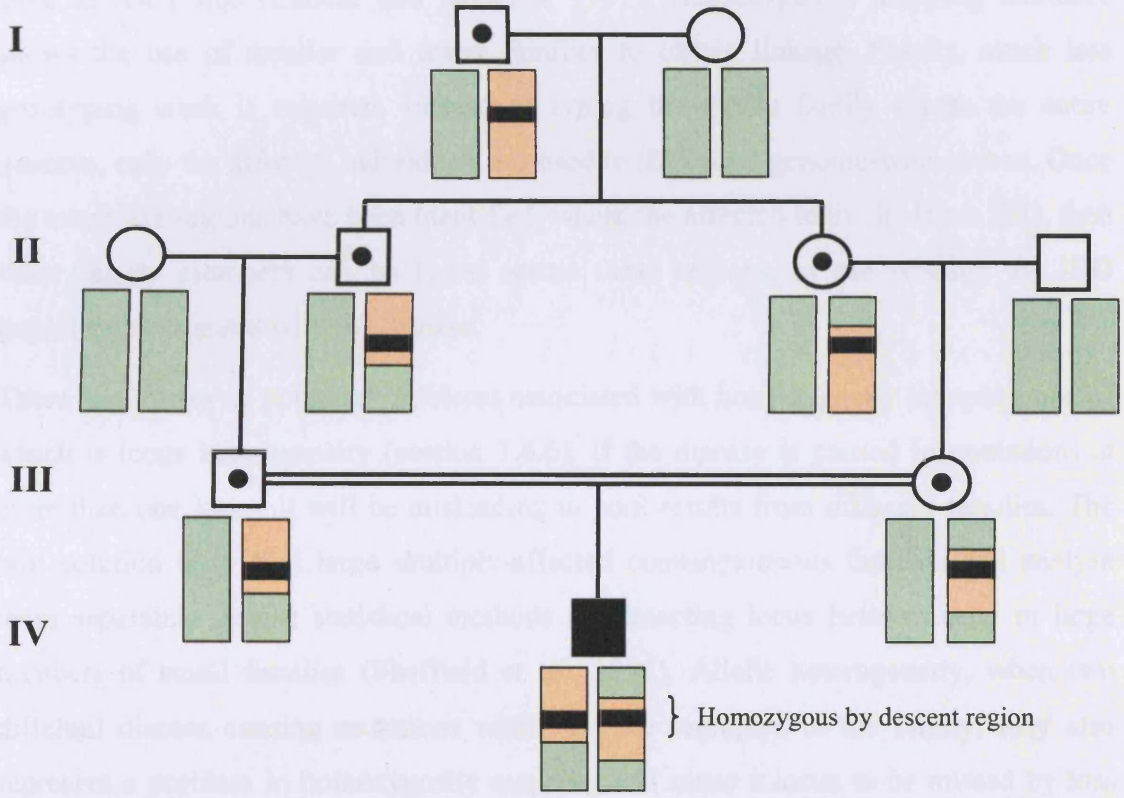


Figure 1.7 Schematic representation of the principle of homozygosity mapping. A simple three-generation first cousin pedigree for an autosomal recessive condition is illustrated. The disease locus is indicated in black and the preserved ancestral area surrounding the locus is shown in orange. Recombination events occurring in generations II and III have narrowed down the region homozygous by descent in the patient. The identification of common regions homozygous by descent between unrelated affected individuals is the basis of homozygosity mapping. The solid symbol represents the affected individual in generation IV. Dotted symbols represent carriers.

and a single affected child are informative for linkage. This is a major advantage when trying to map rare recessive conditions, as families with more than one affected child are often hard to find. Secondly, the approach has high statistical power. A consanguineous family consisting of a single affected child born to first-cousin parents can provide the same support for linkage as a non-consanguineous nuclear family with three affected sibs (Lander and Botstein, 1987). Homozygosity mapping therefore allows the use of smaller and fewer families to obtain linkage. Finally, much less genotyping work is required. Instead of typing the whole family across the entire genome, only the affected individuals are used in the initial genome-wide screen. Once the candidate regions have been identified, where the affected individuals are IBD, then other family members can be typed across these regions, to see whether the IBD haplotype segregates with the disease.

There are, however, potential problems associated with homozygosity mapping, one of which is locus heterogeneity (section 1.6.6). If the disease is caused by mutations at more than one locus, it will be misleading to pool results from different families. The best solution is to find large multiply-affected consanguineous families and analyse them separately or use statistical methods for detecting locus heterogeneity in large numbers of small families (Sheffield *et al.*, 1995). Allelic heterogeneity, when two different disease causing mutations within a gene segregate in the family, may also represent a problem in homozygosity mapping and cause a locus to be missed by loss of homozygosity in flanking markers (Miano *et al.*, 2000). In addition, the identification of a homozygous IBD region unrelated to the disease locus may occur merely by chance (Miano *et al.*, 2000). Finally, the potential for inflation of LOD scores as a result of underestimation of the extent of inbreeding may also be quite common (Miano *et al.*, 2000).

The availability of suitable markers is an important issue. The approach depends on having a good genetic map of the human genome and suitable highly-polymorphic closely-spaced markers covering all parts of the genome evenly.

The choice of marker spacing in the initial genome-wide scan is another critical issue. If markers are too far apart, the locus of interest can be missed, as affected individuals may not be homozygous at the nearest marker to the locus that is tested. But if the

markers are too close together, it will involve testing an excessive number of markers (Genin *et al.*, 1998).

Finally, the degree of consanguinity should be taken into account. The more remote the consanguinity between the parents, the smaller the region of homozygosity around the disease locus in the affected child, so more densely-spaced markers to detect it will be necessary. However, in such cases the interval containing the disease locus can be more precisely mapped.

Over the last 10 years, many genes underlying rare recessive disorders have been mapped using homozygosity mapping. Some of the earliest were genes for alkaptonuria (Pollak *et al.*, 1993), Friedreich ataxia with selective Vitamin E deficiency (Ben Hamida *et al.*, 1993) and Bardet-Biedl (Kwitek-Black *et al.*, 1993a). More recent examples include many genes for autosomal recessive non-syndromic deafness (Kalatzis and Petit, 1998; Willems, 2000); Robinow syndrome (Afzal *et al.*, 2000b; van Bokhoven *et al.*, 2000); recessive hypercholesterolemia (Eden *et al.*, 2001); and Fraser syndrome (McGregor *et al.*, 2003).

1.6.5 Haplotype analysis

The refinement of a candidate region identified by linkage analysis is often achieved by the ascertainment of additional family members and construction of haplotypes using additional markers in the region. The presence of recombinants, especially if they occur in an affected individual, can help to narrow down the region of interest. Recruitment of additional families may also help to refine the candidate region by identifying overlapping regions of homozygosity.

1.6.6 Linkage disequilibrium

Linkage disequilibrium (LDE), or allelic association, can provide a powerful tool to refine the candidate region once the general location of the gene has been established

by conventional linkage analysis. LDE is the association of a particular allele of a polymorphic marker with the disease at a population level. It occurs when apparently unrelated affected individuals all derive their disease chromosome from a common ancestor, and tends to occur in diseases with low mutation rates. This phenomenon is usually only seen with markers closely linked to the disease locus. LDE was successfully employed in the identification of the disease gene for cystic fibrosis (Kerem *et al.*, 1989).

1.6.7 Genetic heterogeneity

Mutations in several genes may produce the same disease phenotype. This is known as locus heterogeneity. Allelic heterogeneity describes the case when different mutations within the same gene cause the same clinical phenotype. The situation in which mutations in the same gene cause different diseases is known as clinical heterogeneity.

Locus heterogeneity represents a complicating factor when the homozygosity mapping approach is adopted as a tool for locating disease genes in a small number of consanguineous families. This problem can be solved by using larger numbers of small consanguineous families and statistical methods which detect heterogeneity or by the use of isolated populations to identify extended inbred kindreds(Sheffield *et al.*, 1995).

For example, three extended Bedouin kindreds were used to map three Bardet-Biedl syndrome loci to chromosome 16q21 (*BBS2*), 3q12-q12 (*BBS*3) and 15q22.3-q23 (*BBS4*) (Kwitek-Black *et al.*, 1993b; Sheffield *et al.*, 1994; Carmi *et al.*, 1995). An extended Bedouin family was also used to map a novel type of autosomal recessive infantile nephronophthisis on chromosome 9q22-31 by homozygosity mapping (Haider *et al.*, 1998). Statistical analysis for linkage heterogeneity can be carried out using the HOMOG program package (Terwilliger, 1994).

1.7 Confirmation of a candidate gene

Once a candidate gene has been identified, by whatever approach, it must be analysed to find evidence for association with the disease. The gene is screened for mutations in a panel of DNAs from unrelated affected individuals and controls. If the selected gene turns out to be the disease causing gene, the unrelated patients will show a number of different mutations predicted to have a deleterious effect on gene function, which segregate with the disease and which are not present in controls. Pairs of specific primers are designed to amplify the coding DNA, either from a genomic DNA sample if the intron/exon boundaries are known, or from cDNA produced by RT-PCR from patient mRNA.

Large-scale deletions, duplications and translocations can be detected by cytogenetic techniques such as fluorescence *in situ* hybridisation (FISH) or array-comparative genome hybridization (CGH). Southern blotting may also be useful to detect large insertions or deletions. Methods employed to screen for small-scale (single base) mutations include heteroduplex analysis, single strand conformational polymorphism, denaturing gradient gel electrophoresis and the protein truncation test. Any sequence variations identified with these methods must be checked by direct DNA sequencing to define the precise nucleotide change(s) involved.

The early chemical methods for DNA sequencing of Maxam and Gilbert required that the DNA fragment to be sequenced was labeled at one end usually by adding either a radioactive phosphate to the 5'- or 3'-end or a nucleotide to the 3'-end. This method has largely been superseded by the enzymatic method of Sanger (dideoxy sequencing method), which uses four specific dideoxynucleotides (ddNTPs) to terminate enzymically synthesized copies of a template. A sequencing primer is annealed to a single strand DNA template molecule and a DNA polymerase extends the primers using dNTPs. The extension reaction is split into four and each quarter is terminated separately with one of the four specific ddNTPs, which act as chain terminators as they lack the 3'-OH group on the deoxyribose necessary to extend the growing chain by the polymerase. Instead of ddNTPs, automated DNA sequencing uses four different

fluorescent dye-labeled chain terminators. All four reactions can be performed in one tube and loaded in one gel lane of the electrophoresis gel. A laser at the bottom of the gel reads the sequence of the four different colours causing the dyes to fluoresce at different wavelengths. The sequence data is given as a series of base specific intensity profiles which can be read either by specific computer programs or manually.

1.8 Aims of the project and overview of thesis

The aim of this project was to identify the disease gene or genes causing JATD and CLS. A number of suggestions about possible candidate genes and loci for JATD had been made before the commencement of this project as described in section 1.1.4. It was decided to look in more detail at the *EVC/EVC2* region, the 12p11-p12 region and the three regions potentially syntenic to the interval to which *srt* maps in the JATD families from whom DNA had been obtained by that time. If these three regions were excluded as showing no evidence of linkage to the JATD locus, a full genome-wide screen was to be undertaken adopting the technique of homozygosity mapping. Suitable candidate genes would then be chosen and screened for disease causing mutations. JATD patients analysed in this study and their families are described in Chapter 2. The results of this investigation are described in Chapter 4 and discussed in Chapter 5. Future directions for this work are also presented in Chapter 5.

This project also describes an investigation into the molecular basis of CLS, aiming to identify the disease gene by adopting a position-independent strategy through an animal model, the *ld* mutant mouse, which shows a similar phenotype to CLS patients. Human *FORMIN-1* (*FMN-1*) and *GREMLIN* were identified as candidate genes for the disease and were screened for mutations in a panel of CLS patients. Patients are described in Chapter 2 and results are described and discussed in Chapter 3.

Chapter 2 Patients, Materials and Methods

2.1 Patients

2.1.1 JATD patient samples

The JATD patients and their families studied in this project were ascertained with the approval of the Institute of Child Health Ethics Committee in several ways. Family 1 was recruited through clinicians working in the Clinical and Molecular Genetics Unit at Great Ormond Street Hospital. Other patients were ascertained following calls for patients or direct discussion with clinical geneticists working both in the UK and abroad at the British Society of Human Genetics Conferences in autumn 2001 and 2002; the Dysmorphology Club Meeting of spring 2001; the Manchester Birth Defects Conference in autumn 2002; and the European Society of Human Genetics Conferences (held in spring 2001 and 2003). Further notices explaining the project and calls for patients were also published in the *British Society of Human Genetics Newsletter* and the *American Journal of Human Genetics*. The corresponding authors of previously published papers on JATD families were contacted and asked whether they and the families were interested in participating in the project. The Jeune Syndrome Information and Support Group in Italy was also contacted.

Patients were diagnosed with JATD by the referring clinician through clinical and radiological examinations according to the diagnostic criteria described in section 1.1.1. Where possible, clinical photographs and radiographs were obtained from the referring clinician and reviewed by Prof. Christine Hall to confirm the diagnosis.

For most of the patients and their families, DNA samples or stored blood were already available for laboratory analysis. When this was not the case, the referring clinicians were asked to obtain samples of venous blood from consenting individuals, and forward these for DNA extraction. In the case of one deceased child (Family 7, patient

II.3), paraffin embedded tissue samples were sent for DNA extraction. Additional venous blood samples were obtained from patients IV.1 and IV.3 in Family 1 and from patients VI.3 and VI.5 in Family 2 in order to establish lymphoblastoid cell lines.

For the work described in this thesis, DNA samples, venous blood or tissue samples were obtained from a total of 35 affected individuals from 12 consanguineous and 19 non-consanguineous families. The 12 consanguineous families comprised 9 first cousin parent matings, 1 third cousin mating, 1 multiply consanguineous mating and 1 family in which the parents were distantly related. All pedigrees were consistent with autosomal recessive inheritance.

A collaboration with Dr Colin Johnson, Dr Louise Brueton and Prof. Eammon Maher at the University of Birmingham Medical School was set up in autumn 2001. They had identified five affected consanguineous families, three of which originated in Pakistan but were resident in the UK; the other two originated in Italy and France. They had also obtained DNA samples from six affected non-consanguineous families. At the beginning of 2003, a collaboration with Dr Valerie Cormier-Daire at the Hôpital Necker in Paris was set up. Her group had identified seventeen affected families, ten of which were consanguineous.

For the study described in this thesis, all DNA samples were ascertained by our group except for DNA samples from Families 8, 9 and 16 provided by collaborator Dr Colin Johnson and DNA samples from Family 10 sent by collaborator Dr Valerie Cormier-Daire.

2.1.1.1 Clinical details of consanguineous families

(i) Family 1

This multiply-consanguineous Pakistani family (Fig. 2.1) was clinically assessed by Dr Louise Wilson, Consultant Clinical Geneticist at Great Ormond Street Hospital, and I myself was able to meet them when they were reviewed in the Genetics Clinic. There were two affected boys (IV.1 and IV.3), each the child of a different first-cousin marriage (they are double first cousins). Both had a very mild form of JATD.

Patient IV.3 (Fig. 2.2 and 2.3) was noted on antenatal ultrasound (US) to have short femora. He had no respiratory problems in the neonatal period. He was referred to the Genetics Clinic at GOSH from Thorpe Coombe Children's Centre because of short limbs and a relatively large head. He was first seen at Great Ormond Street Hospital at the age of 14 months. His clinical features were consistent with a diagnosis of JATD. His height was between the 9th and the 25th centiles, while his weight was on the 50th centile. He had mild rhizomelic limb shortening, a small narrow chest, a protuberant abdomen, short broad hands and fingers and small feet. A small number of oral mucosal frenulae were observed. He had a bossed forehead, sparse and patchy scalp hair and sparse eyebrows with normal eyelashes. His developmental progress was normal. A full skeletal survey was performed and reviewed by Prof. Christine Hall, who confirmed the diagnosis of JATD (Fig. 2.3). Chromosome studies, renal assessment and eye examination were normal. At the age of $4^1/_2$ years, his height remained between the 9th and 25th centiles. To date he has not developed any extra-skeletal problems, but he continues to be followed up in the renal and eye clinic.

Both his parents (III.2 and III.3) and his older sister (IV.2) were unaffected, with normal stature. His mother (III.3) had serial US scans during her next pregnancy, but no limb abnormalities were detected. When the baby (IV.4) was born, he too was found to be unaffected. His full skeletal survey was normal.

Family 1

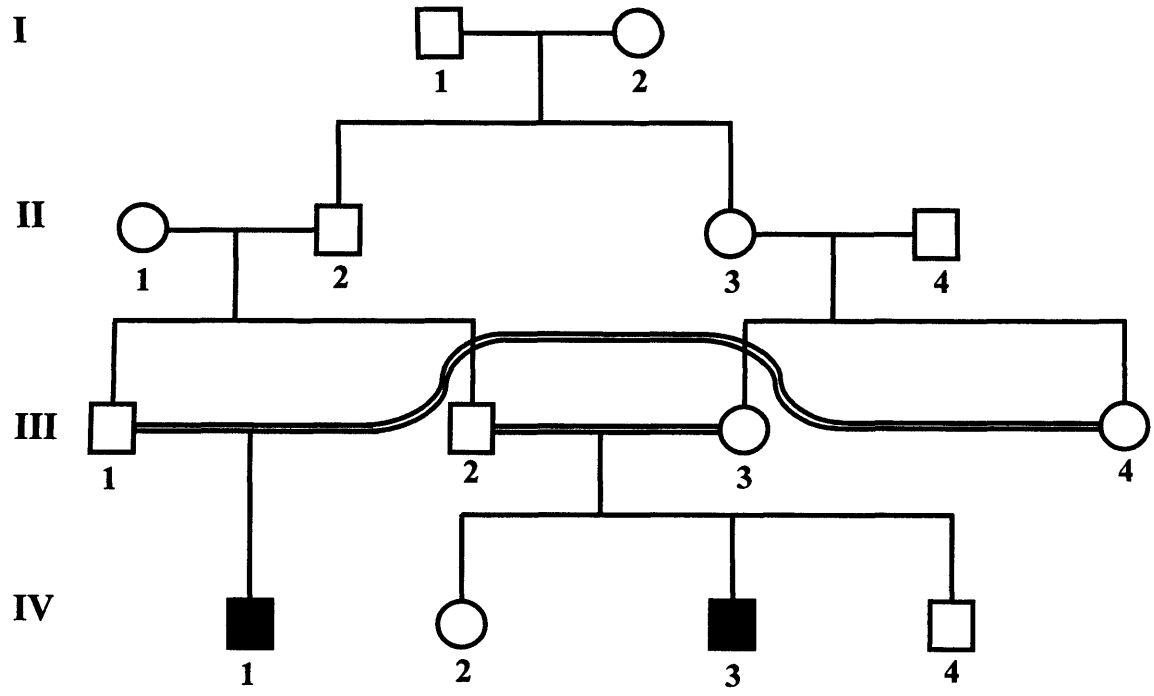


Figure 2.1 Pedigree drawing of Family 1, a multiply-consanguineous family with Jeune syndrome. Solid symbols represent affected individuals.

Figure 2.2 Clinical photographs of patient IV.3 from Family 1. He has mild rhizomelic limb shortening and a narrow chest (A), with a bossed forehead, sparse scalp hair (B) and small broad hands with short fingers (C).

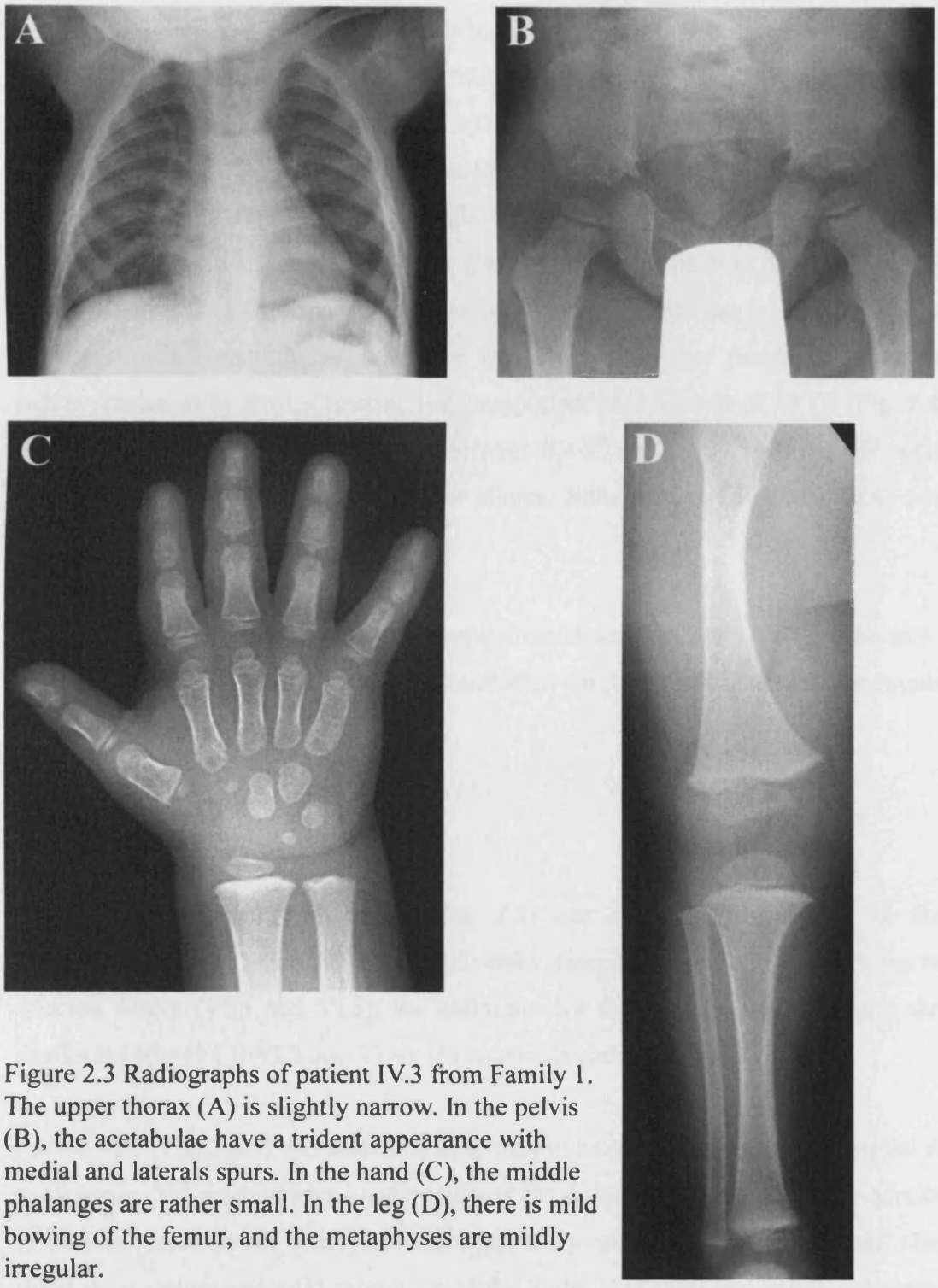


Figure 2.3 Radiographs of patient IV.3 from Family 1. The upper thorax (A) is slightly narrow. In the pelvis (B), the acetabulae have a trident appearance with medial and laterals spurs. In the hand (C), the middle phalanges are rather small. In the leg (D), there is mild bowing of the femur, and the metaphyses are mildly irregular.

Patient IV.1 (Fig. 2.4) was also noted to have short femora on antenatal US. He too had no respiratory problems in the neonatal period. He was referred to the Genetics Clinic at Great Ormond Street Hospital from Whipps Cross Hospital because of short limbs and was seen at Great Ormond Street Hospital at the age of 12 months. His main clinical features, again consistent with JATD, were mild rhizomelic limb shortening, a small narrow chest, a protuberant abdomen, broad hands and 5th finger clinodactyly. His height was on the 50th centile, while his weight was between the 75th and 91st centiles. He also had a few lower jaw gum frenulae and he was born with a right inguinal hernia and hypospadias (neither of which are particular features of JATD). His development was normal, and there were no renal or eye problems. A skeletal survey, reviewed by Prof. Christine Hall, supported the diagnosis of JATD (Fig. 2.4). At the age of $3^3/_4$ years, his height was between the 9th and the 25th centiles. He is also being followed up in the renal and eye clinics. Both parents (III.1 and III.4) were unaffected, and he has no siblings.

DNA samples were obtained from affected individuals (IV.3 and IV.1), both sets of parents (III.2 and III.3; III.1 and III.4), unaffected sib (IV.2) and maternal grandmother (II.3).

(ii) Family 2

This consanguineous Dutch family (Fig. 2.5) was clinically assessed by Dr Joep Tuerlings, Clinical Geneticist at the University Hospital Nijmegen. There were two affected sisters (VI.3 and VI.5), the children of a third-cousin marriage, and three unaffected sibs (VI.1, VI.2 and VI.4). Both patients had a mild form of JATD.

Patient VI.3 (Fig. 2.6A) was delivered at home, but had to be transferred to hospital due to asphyxia. She was intubated and ventilated for about six weeks and was re-admitted to hospital twice in her first year with respiratory problems. She has a small chest, slight short stature and mild shortening of the limbs. Her chest and pelvic radiographs are said to be typical of JATD. Her renal function and eyesight are being monitored and

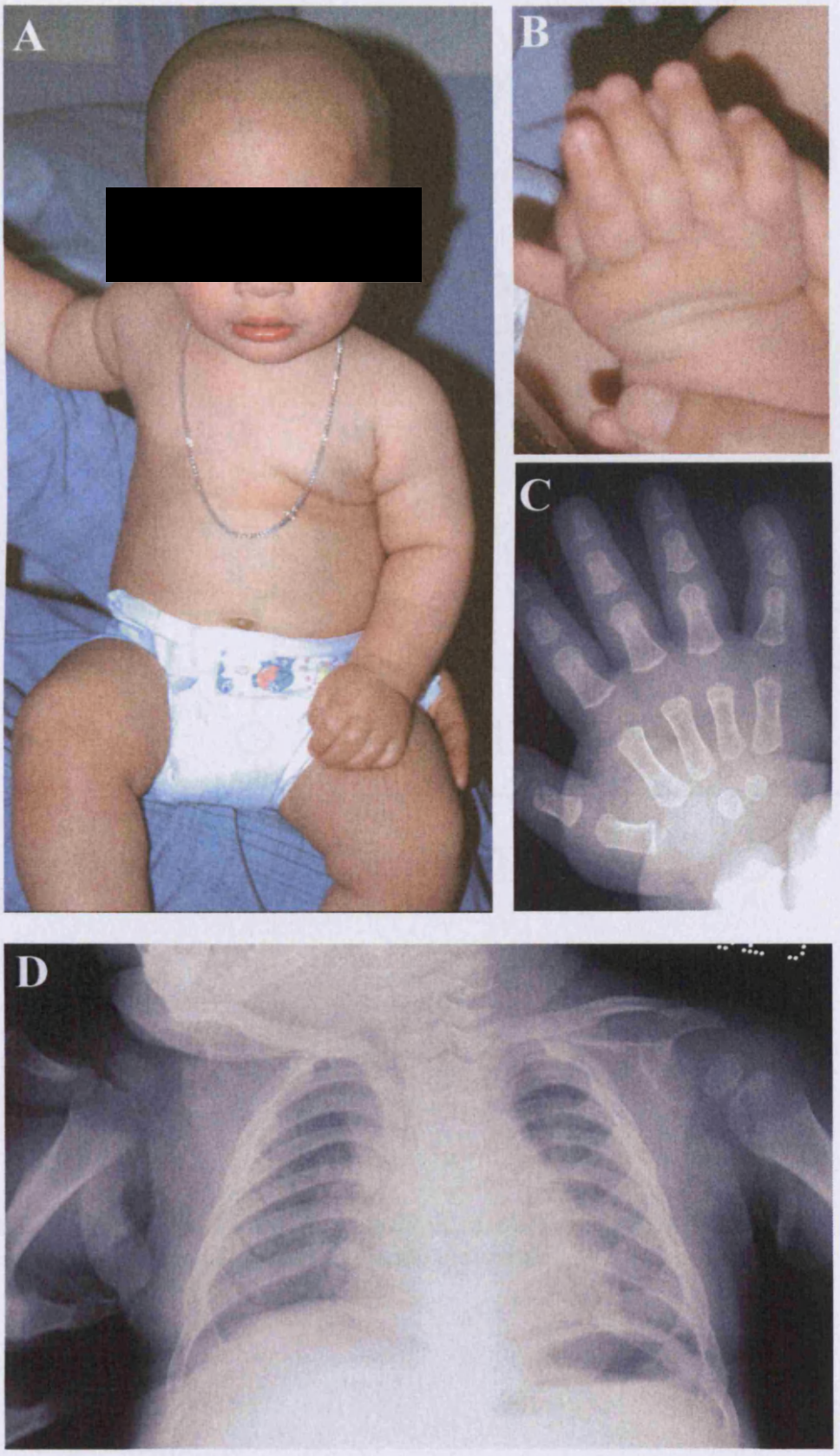


Figure 2.4 Clinical photographs and radiographs of patient IV.1 from Family 1. There is mild rhizomelic limb shortening (A) and the hands are broad (B) with 5th finger clinodactyly (C). The upper thorax is rather narrow (D).

Family 2

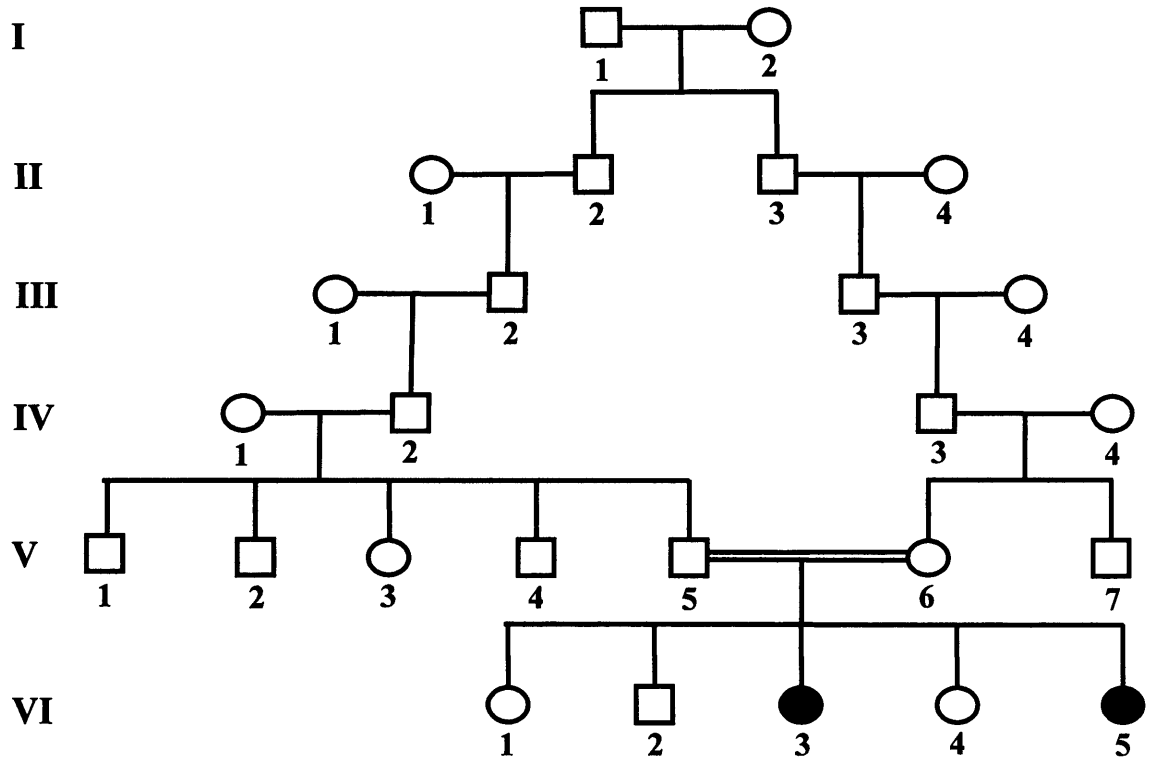


Figure 2.5 Pedigree drawing of Family 2, a consanguineous family with Jeune syndrome. Solid symbols represent affected individuals.

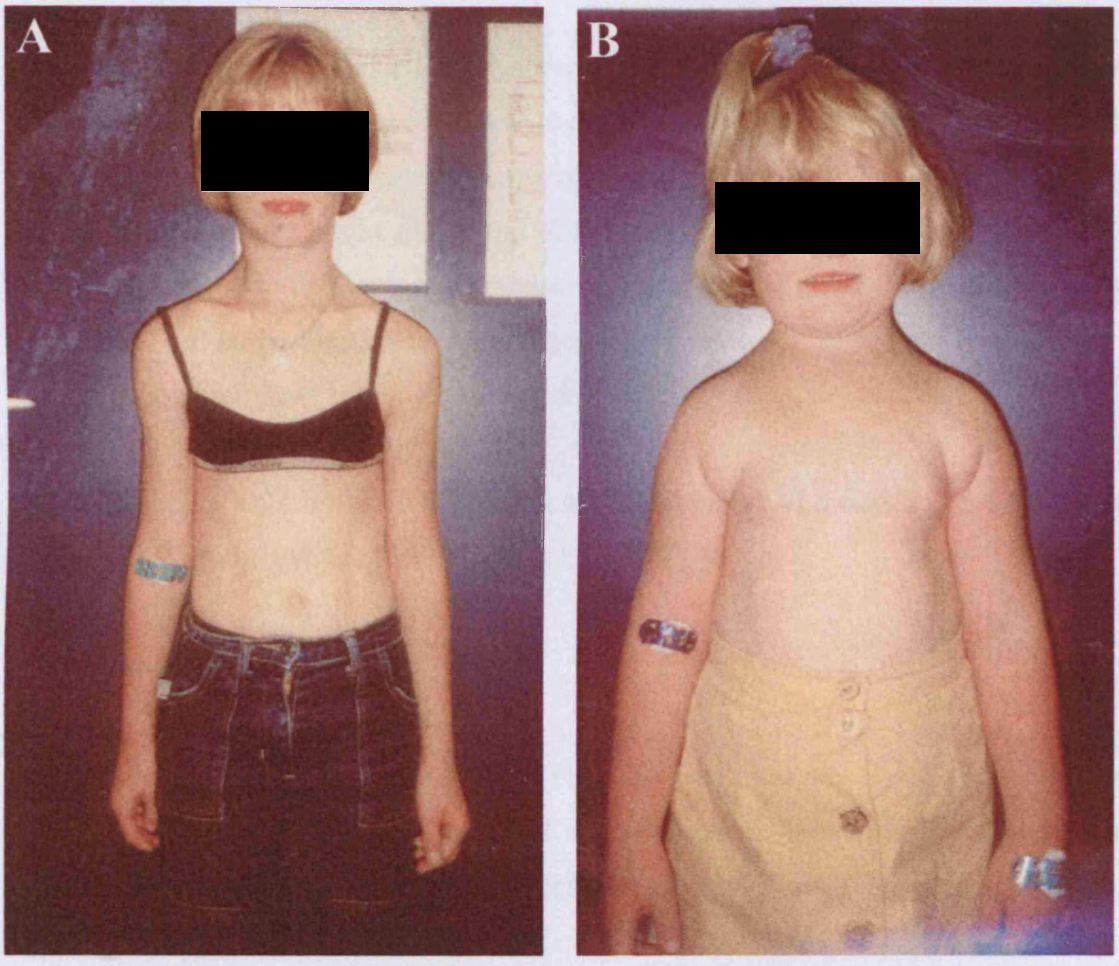


Figure 2.6 Clinical photographs of patients VI.3 (A) and VI.5 (B) from Family 2. Both patients have a narrow chest and mild rhizomelic limb shortening.

so far no abnormalities have arisen. She has also feeding difficulties and speech problems, and, at the age of 17 years, now attends a special school.

Patient VI.5 (Fig. 2.6B and 2.7) also suffered from asphyxia at birth. She was intubated and ventilated for about two weeks, re-admitted three times in her first year with respiratory problems and required home oxygen for a period. Like her sister, she has a small chest and only slight shortening of the limbs. She had changes typical of JATD on her chest, pelvic and limb radiographs, which have been reviewed by Prof. Christine Hall. She had no problems with renal function and eyesight. Now, aged 10 years, she is making good progress, although she has developed asthma.

DNA samples were obtained from affected individuals (VI.3 and VI.5), both parents (V.5 and V.6) and two unaffected sibs (VI.2 and VI.4).

(iii) Family 3

This consanguineous Pakistani family (Fig. 2.8) was clinically assessed by Dr Jenny Morton, Consultant Clinical Geneticist at Birmingham Women's Hospital and has recently been reported in the literature (described as proband A)(Morgan *et al.*, 2003).

There was only one affected individual (II.1), the child of unaffected first cousin parents. She was 39 months of age at the time of the study and had relatively mild JATD with short stature, short upper and lower limbs and a very narrow chest. She had respiratory problems in the neonatal period and was found to have rod/cone dystrophy at 9 months, but had normal renal and hepatic function by 39 months.

A full skeletal survey was performed and reviewed by Prof. Christine Hall, who confirmed the diagnosis of JATD. Her thorax was narrow and small with short horizontal ribs. She had handle-bar shaped clavicles, shortening and bowing of the radius and ulna with subluxation of the radial head, small iliac bones and large rounded capital femoral epiphyses.

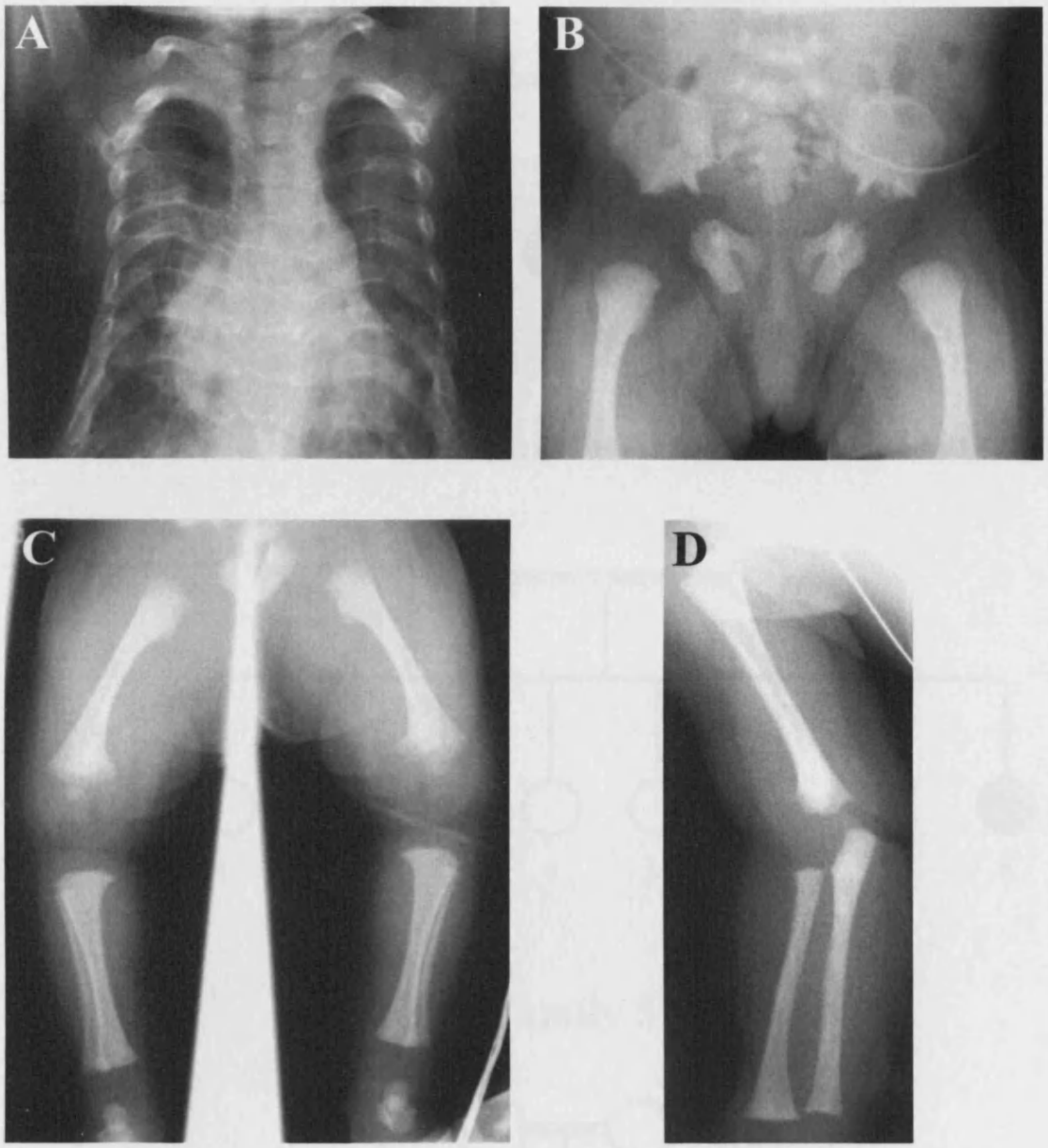


Figure 2.7 Radiographs of patient VI.5 from Family 2. There is significant shortening of the ribs (A), a trident appearance of the acetabulae (B), mild shortening of the humerus (C) and mild shortening and bowing of the femora with few metaphyseal spurs (D).

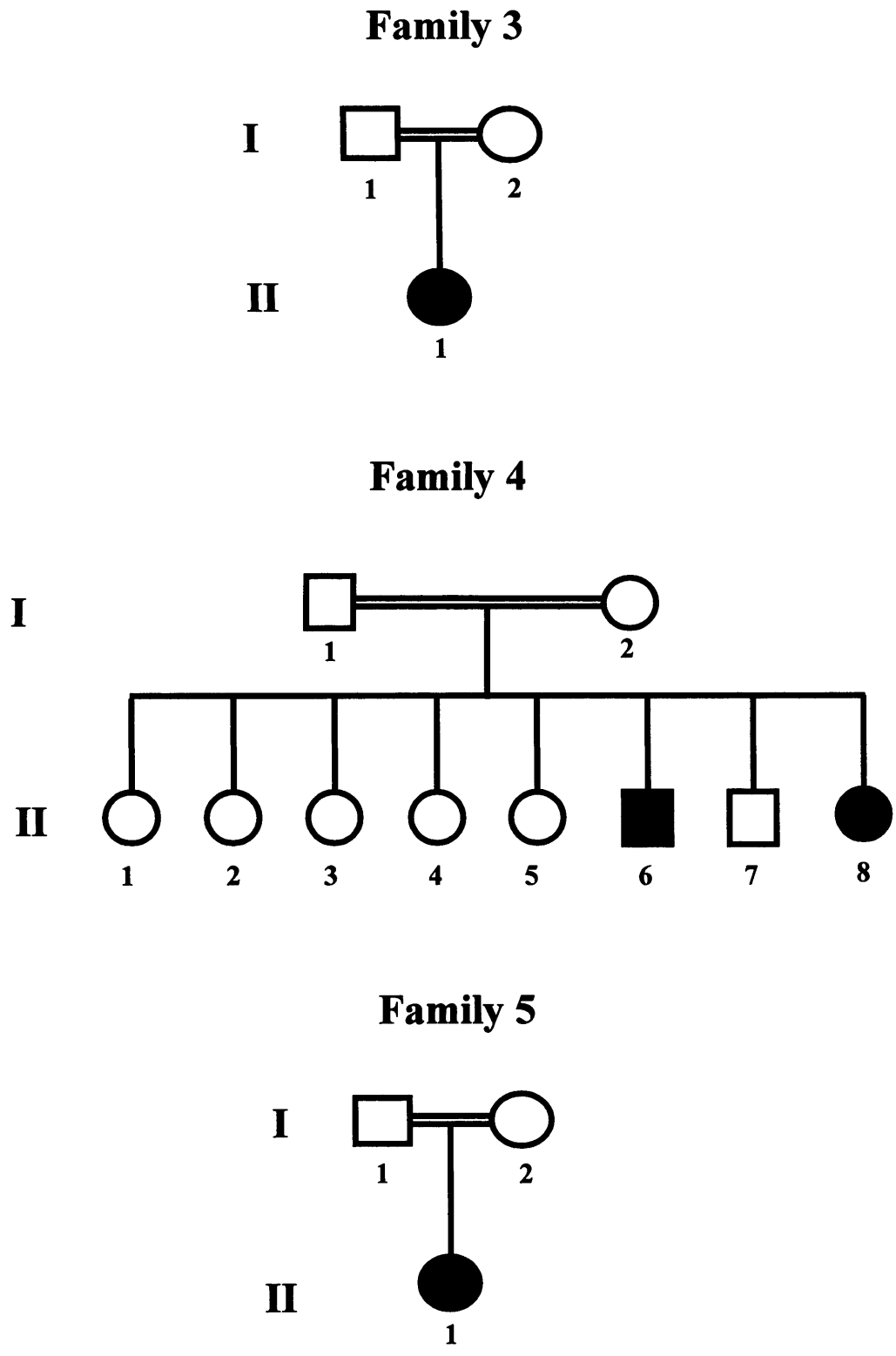


Figure 2.8 Pedigree drawing of Families 3, 4 and 5, three consanguineous families with JATD. Solid symbols represent affected individuals.

DNA samples from the proband (II.1) and her parents (I.1 and I.2) were sent to our laboratory by Prof. Judith Goodship, Consultant Clinical Geneticist in Newcastle. Our collaborators in Birmingham had also independently obtained DNA from this family.

(iv) Family 4

This consanguineous Arab family (Fig. 2.8) was clinically assessed by Dr Lihadh Al-Gazali, Consultant Clinical Geneticist at the Faculty of Medicine and Health, Al Ain, United Arab Emirates. The proband (II.6) was the sixth child of first cousin parents. There were six unaffected sibs, five older sisters (II.1-II.5) and one younger brother (II.7).

The patient (Fig. 2.9 and 2.10) had short stature, short limbs, a narrow chest and postaxial polydactyly of the right foot. He had an operation to remove the extra-digit and the fifth and the sixth toes were found to share the same metatarsal. He had normal fingernails but dysplastic toenails, which is not a feature of JATD but of EVC. He had no neonatal respiratory problems but had recurrent mild chest infections from age 1½ years to the present, a chronic cough and airway hyper-reactivity. He had a systolic heart murmur, but his echocardiogram was normal. Renal US soon after birth showed possible mild dilatation of the left renal pelvis, but repeat US at 1½ years of age was normal. He had no dysmorphic features (apart from a depressed nasal bridge which is familial) and his development was normal. Now, aged 5 ½ years, he has no retinal, renal or hepatic problems. Radiographs were reviewed by Prof. Christine Hall who confirmed the diagnosis of JATD (Fig. 2.10).

Very recently another affected baby (II.8) was born in this family. She is said to have a very small chest, bilateral postaxial polydactyly in the hands and no nail dysplasia. Radiographs are said to be typical of JATD.

Initially DNA samples from the proband (II.6) and his parents (I.1 and I.2) were obtained from Dr Judith Goodship. DNA samples were subsequently obtained from the

Figure 2.9 Clinical photographs and pre-operative radiograph of the feet of patient II.6 from Family 4. He has short stature, a narrow chest (A) and dysplastic toenails (B). The right 5th and 6th toes share the same metatarsals (C).

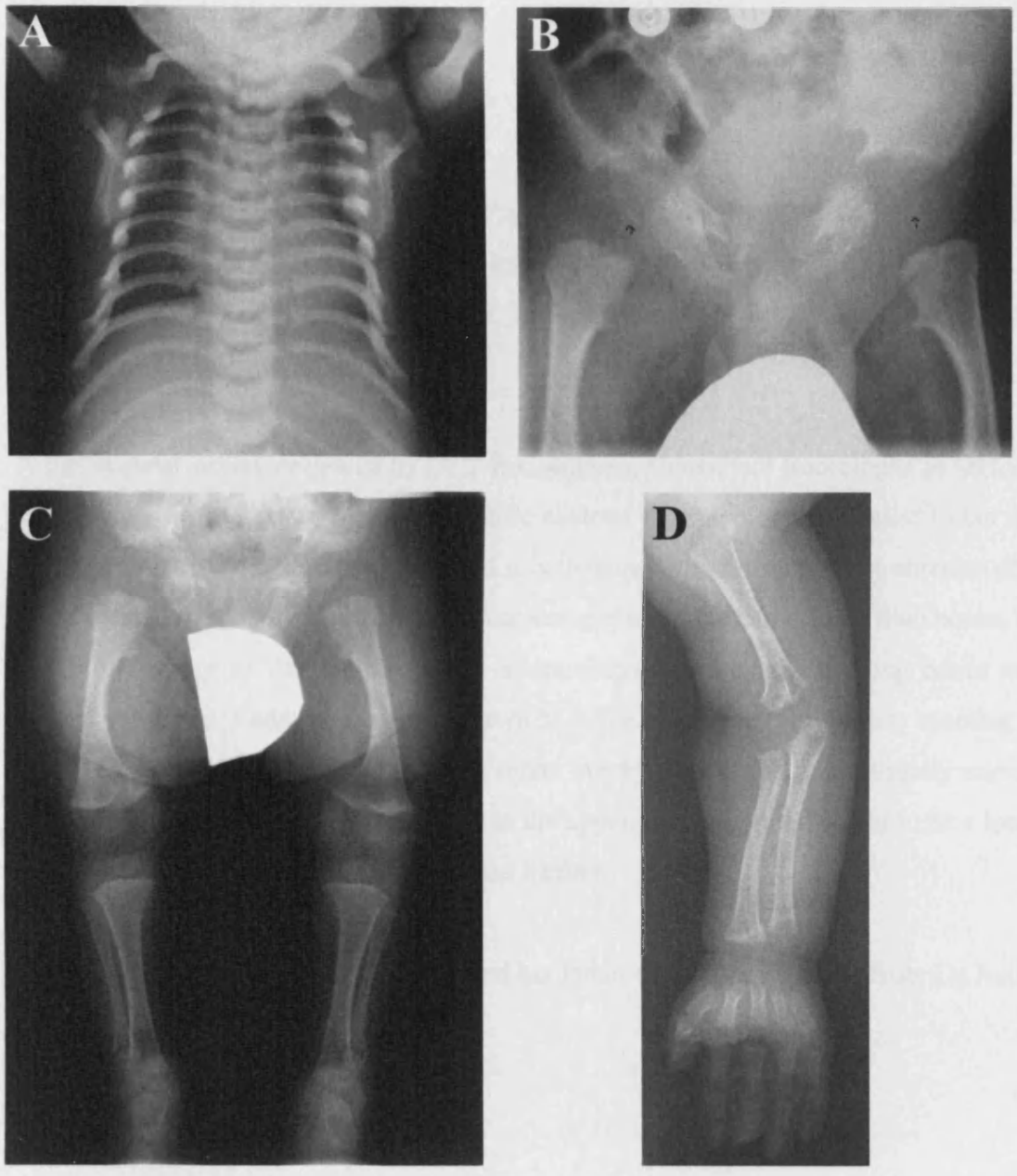


Figure 2.10 Radiographs of patient II.6 from Family 4. The upper thorax is narrow with short ribs (A). In the pelvis, the acetabulae have a trident appearance with medial and lateral spurs (B).There is shortening of the long bones, bowing of the femora, disproportionate shortening of the fibulae and advanced ossification of the capital femoral epiphyses (C). In the hand, there are coned shaped epiphyses of the middle phalanges (D).

six unaffected sibs and very recently from the affected baby (II.8) directly from Dr Lihadh Al-Gazali.

(v) Family 5

This consanguineous Pakistani family (Fig. 2.8) was clinically assessed by Dr Richard Gibbons and Dr Ed Blair, Consultant Clinical Geneticists at Oxford. The proband (II.1) was the only child of first cousin parents. He had short stature and signs of mild spastic diplegia.

A full skeletal survey reviewed by Dr T R Goodman, Consultant Radiologist in Oxford, showed that he had an unusual non-specific skeletal dysplasia that is similar to but not typical of either JATD or EVC. He had a bell-shaped thorax with short anterior ribs, some bowing of the left femur and mild lumbar scoliosis. He had normal iliac bones, no trident deformity of the acetabulae, no shortening/thickening of the long bones and normal vertebrae. Radiographs were shown at a Skeletal Dysplasia Society meeting in 2002 and no definite diagnosis could be made. An MRI scan showed a slightly narrow spinal canal with an area of abnormality in the upper cervical cord thought to be a long-standing lesion which was not investigated further.

DNA samples from the proband (II.1) and his father (I.1) were obtained from Dr Judith Goodship.

(vi) Family 6

This consanguineous Turkish family (Fig. 2.11) was clinically assessed by Dr Beyhan Tuysuz, Consultant Clinical Geneticist at the University of Istanbul, Turkey. The proband (II.2) was the second child of first cousin parents. She had short stature, a small chest and postaxial polydactyly in the feet. She had no respiratory, renal or retinal complications. Radiographs were reviewed by Prof. Christine Hall who confirmed the

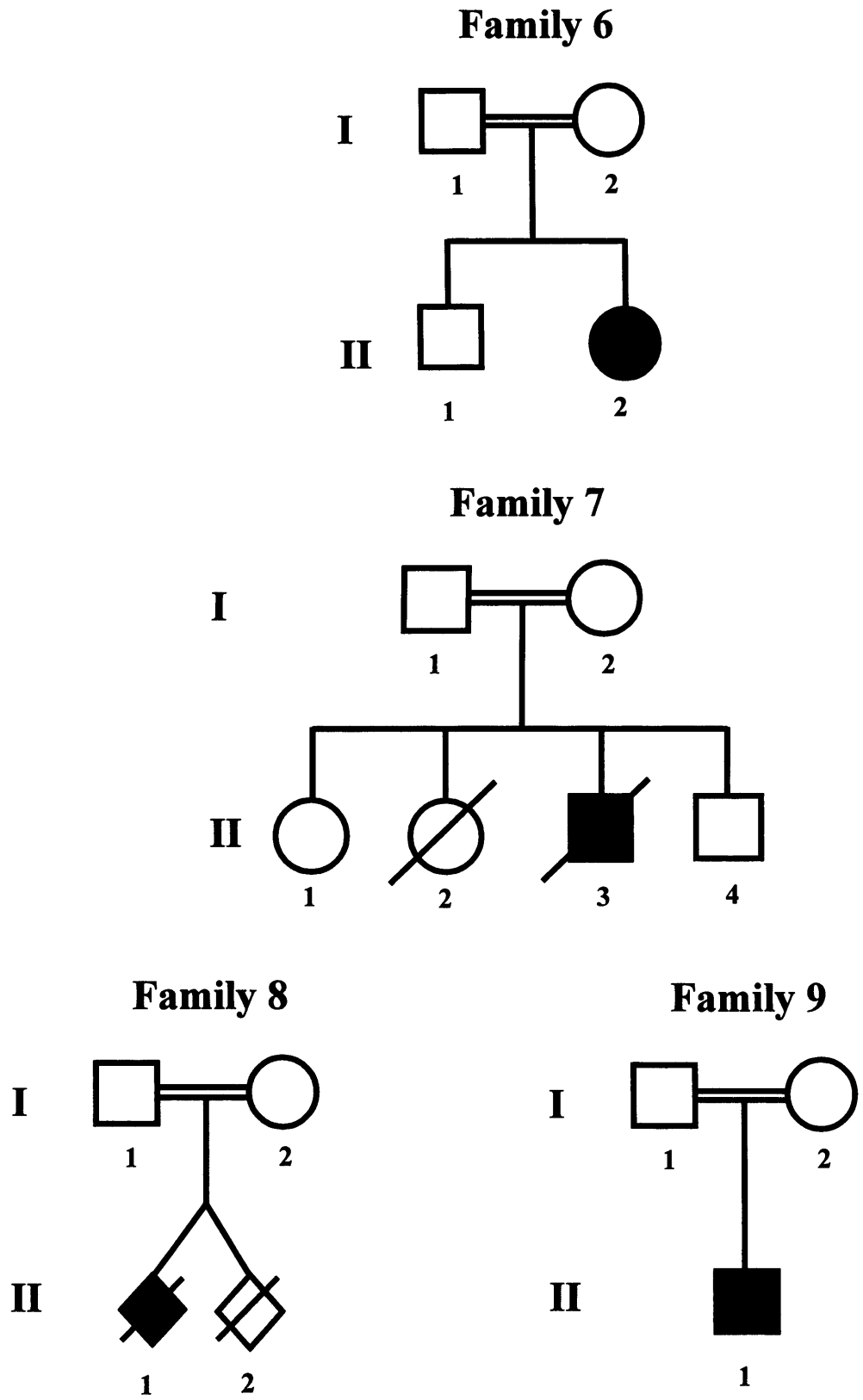


Figure 2.11 Pedigree drawing of Families 6, 7, 8 and 9, four consanguineous families with JATD. Solid symbols represent affected individuals.

diagnosis of JATD (Fig 2.12). DNA samples were obtained from all members of the family (I.1, I.2, II.1 and II.2).

(vii) Family 7

This consanguineous Turkish family (Fig. 2.11) was also clinically assessed by Dr Beyhan Tuysuz. There was one deceased affected child (II.3), one unaffected sister (II.1), one unaffected brother (II.4) and one deceased unaffected sister (II.2), all children of first cousin parents. The proband (II.3) had short stature and a very small chest and died at 3½ months because of respiratory insufficiency. Diagnosis of JATD was confirmed by Prof. Christine Hall who reviewed his radiographs (Fig. 2.13A). He had a narrow thorax, trident appearance of the acetabulae and advanced ossification of the upper femoral epiphyses.

DNA samples were obtained from the unaffected sister (II.1) and brother (II.4) and both parents (I.1 and I.2) as well as paraffin embedded tissues from the proband (II.3).

(viii) Family 8

This consanguineous Pakistani family (Fig. 2.11) was clinically assessed by Dr Jill Clayton-Smith, Consultant Clinical Geneticist at St Mary's Hospital, Manchester and DNA samples were sent to our collaborator Dr Colin Johnson in Birmingham. The case has recently been reported in the literature (described as proband C)(Morgan *et al.*, 2003). The parents were first cousins and they had dizygotic twin fetuses, one severely affected (II.1) and one unaffected (II.2).

Fetus II.1 (Fig. 2.13B) was stillborn at 30 weeks and had a narrow chest. There were no detectable abnormalities in the liver, kidney or heart on post-mortem examination. Prof. Christine Hall reviewed the fetal radiographs and confirmed the diagnosis. He had the typical skeletal abnormalities of JATD in the thorax and pelvis.

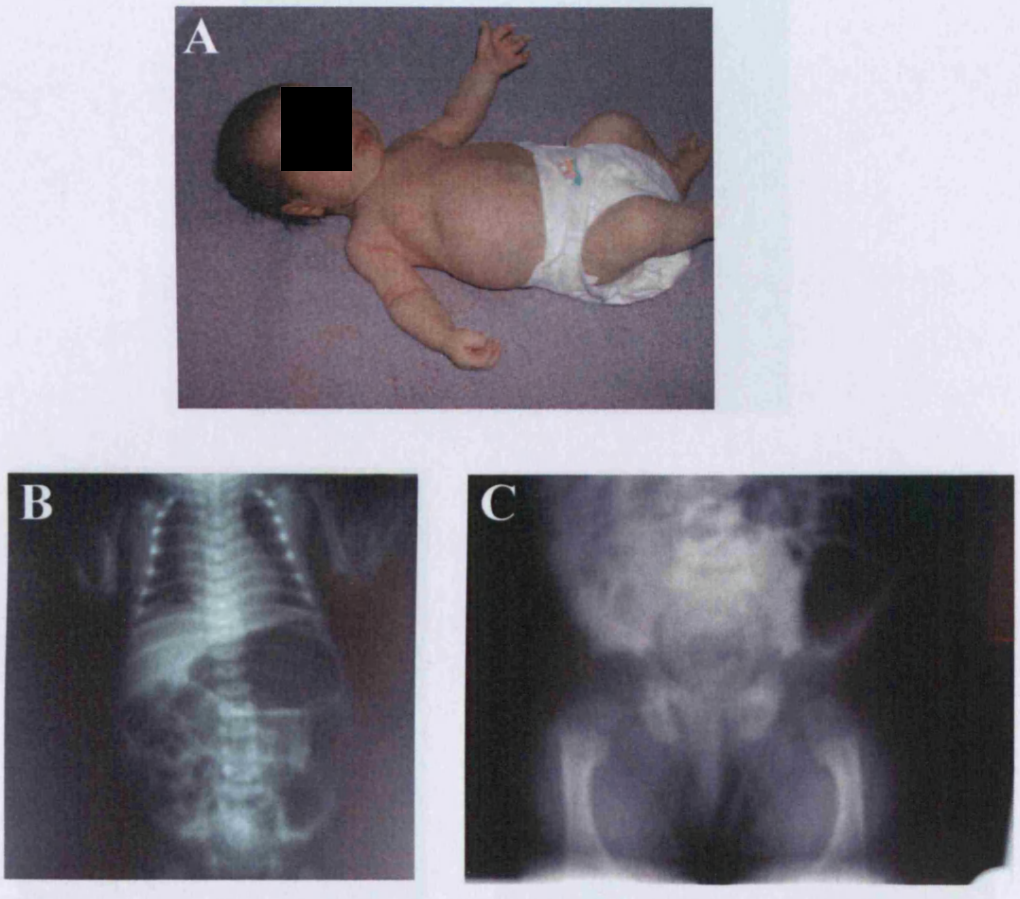


Figure 2.12 Clinical photograph and radiographs of patient II.2 from Family 6. She has a narrow chest (A) with short ribs (B). The acetabulae have a trident appearance and the iliac bones are rather short (C). There is advanced ossification of the upper femoral epiphyses and metaphyseal irregularities (C).

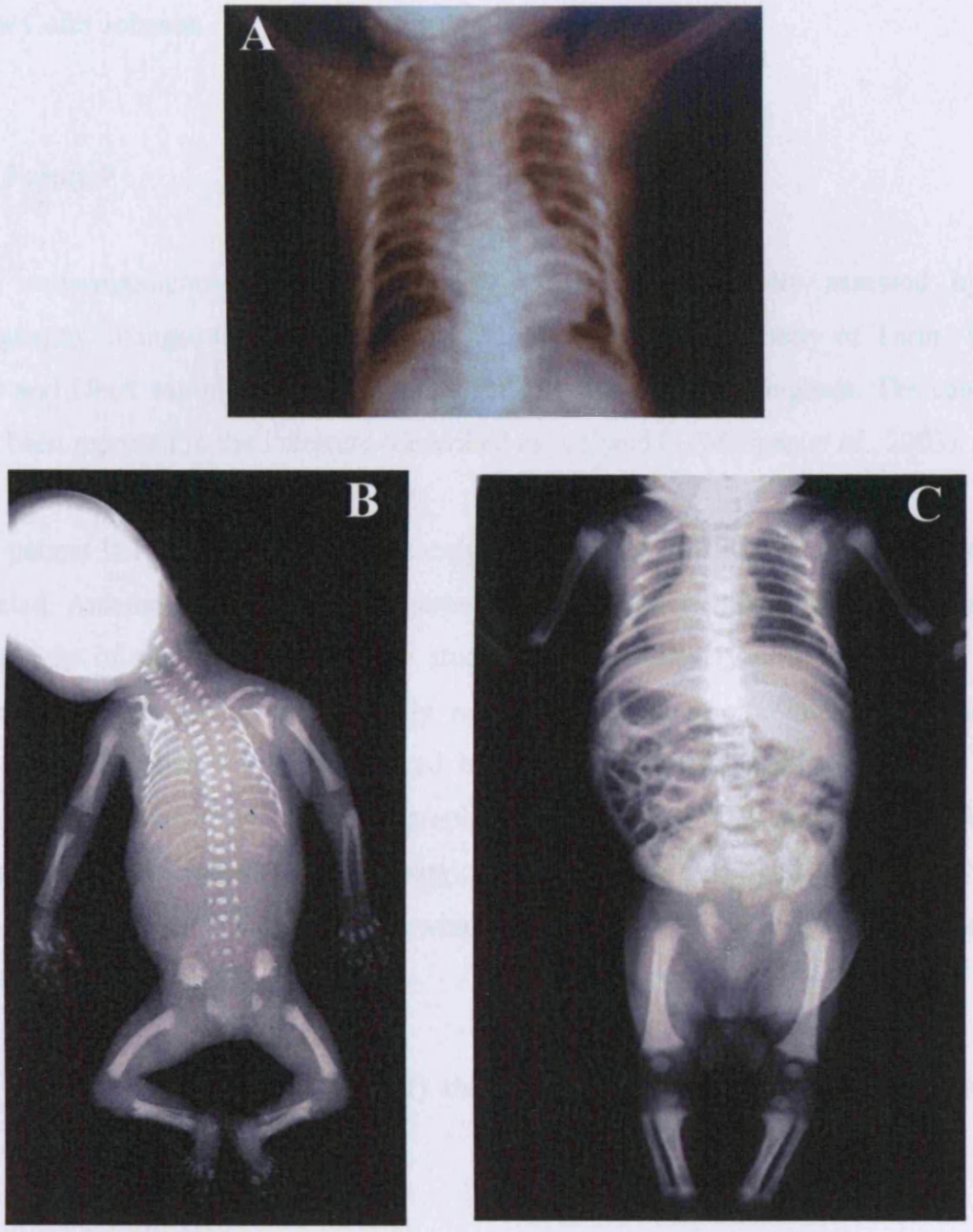


Figure 2.13 Radiographs of patients II.3 from Family 7 (A), II.1 from Family 8 (B) and II.1 from Family 9 (C). Patients have a small thorax with short ribs. Proband II.1 from Family 8 (B) has medial and lateral spurs on the acetabular roofs and additional lateral spurs on the iliac wings. Mesomelic shortening of the lower limbs with some cupping of the metaphyses is also present. Femurs are mildly curved. In the hands, there is shortening of the middle phalanges and metacarpals, but no polydactyly. Proband II.1 from Family 9 (C) has mild narrowing of the thorax and horizontal acetabular roofs with spur-like projections. There is rather severe mesomelic limb shortening and significant irregularity of the distal ends of the femora with flared metaphyses.

DNA samples from the affected fetus (II.1) and his parents (I.1 and I.2) were provided by Dr Colin Johnson.

(ix) Family 9

This consanguineous Italian family (Fig. 2.11) was clinically assessed by Dr Margherita Silengo, Consultant Clinical Geneticist at the University of Turin, Turin, Italy and DNA samples were sent to Dr Colin Johnson in Birmingham. The case has also been reported in the literature (described as proband D)(Morgan *et al.*, 2003).

The patient II.1 (Fig. 2.13C) was the only child of first cousin parents and was mildly affected. Antenatal US at 24 weeks gestation revealed shortening of the limbs. He was 30 months of age at the time of the study and had relatively mild JATD with short upper and lower limbs and a mildly narrow chest (chest circumference below 3rd percentile). He had normal renal and hepatic function and normal eyesight. Prof. Christine Hall reviewed the radiographs and confirmed the diagnosis. However, because of the extremely mild radiological findings, she described the child as not a classic JATD case. He had mild narrowing of the thorax and some minor changes of the acetabular roofs (see Fig. 2.13C).

DNA samples from the proband (II.1) and both parents (I.1 and I.2) were provided by Dr Colin Johnson.

(x) Family 10

This consanguineous Moroccan family (Fig. 2.14) was clinically assessed by Dr A. Toutain, Consultant Clinical Geneticist at Tours, France. The proband (II.10) was the 10th child of first cousin parents. There were six affected deceased sibs and three unaffected sibs. The proband (Fig. 2.15A), now 8 years old, had the typical features of JATD such as short stature, short limbs, a narrow chest and postaxial polydactyly. She

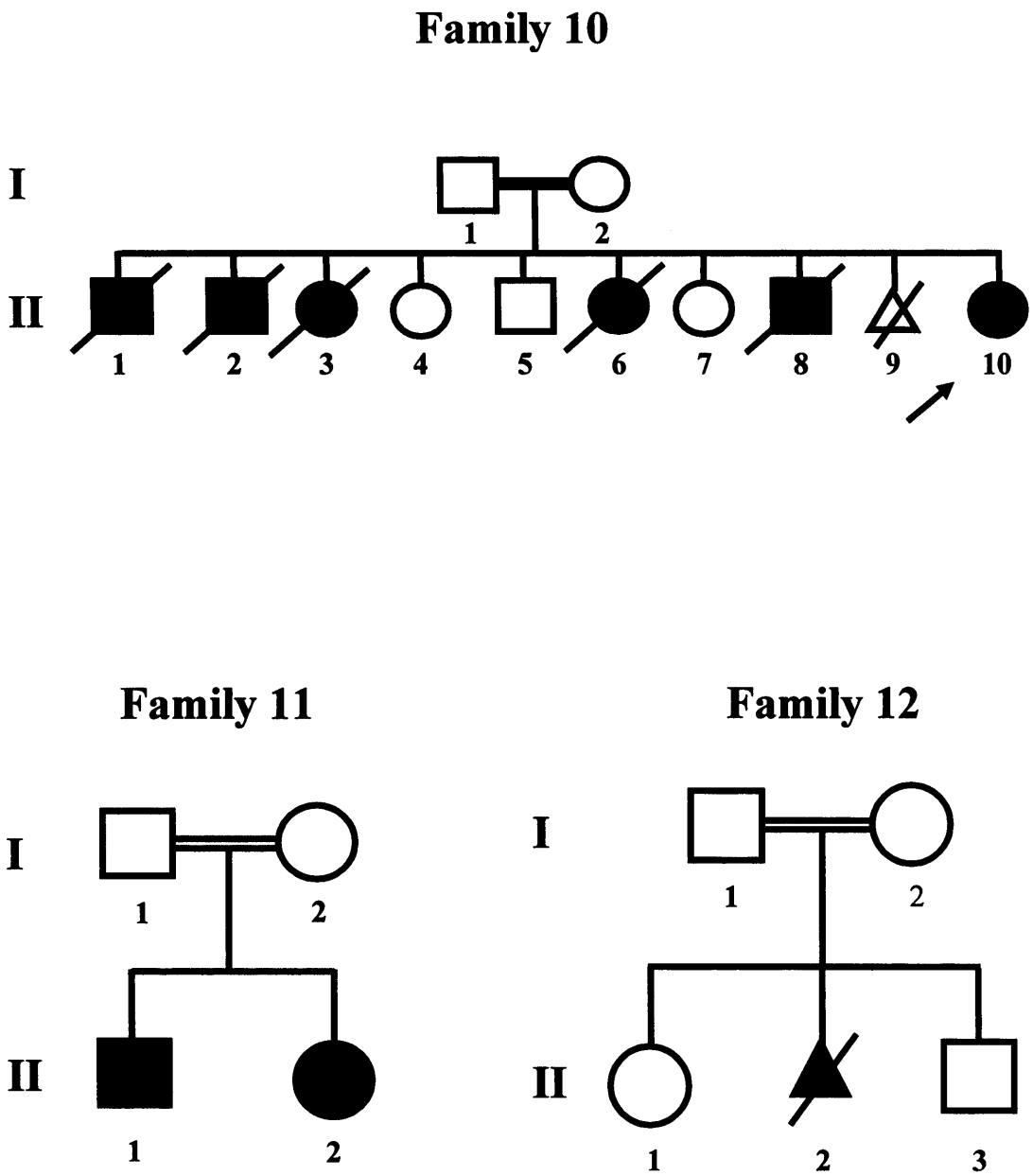


Figure 2.14 Pedigree drawing of Families 10, 11 and 12, three consanguineous families with JATD. Solid symbols represent affected individuals.

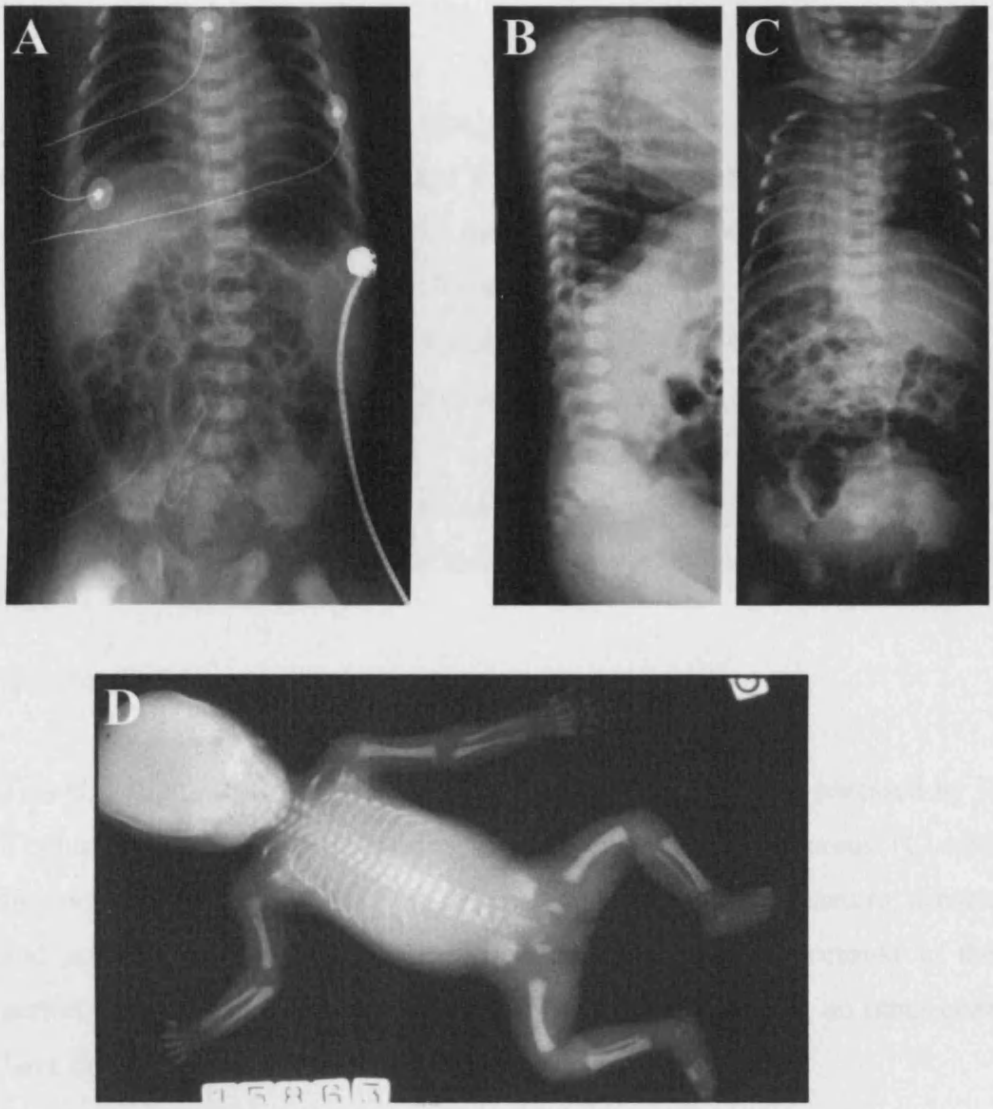


Figure 2.15 Radiographs of patients II.10 (A), II.1 (B and C) and fetus II.9 (D) from Family 10 showing the typical skeletal abnormalities of JATD. They have narrow thoraces with short ribs and the acetabulae have a trident appearance.

also had severe mental retardation, never previously associated with JATD, and Dandy-Walker malformation which has been reported in one other JATD case (Silengo *et al.*, 2000).

From the six affected deceased sibs, II.1 died at age 3 days from respiratory failure (Fig. 2.15B and C), II.3 died at age 2 months also from respiratory failure, II.6 died at age 3 days and had cleft palate, II.8 died at age 24 days from respiratory failure and II.9 was a termination of pregnancy for cerebellar anomaly (Fig. 2.15D). Radiographs of affected individuals II.1, II.3, II.9 and proband II.10 were reviewed by Prof. Christine Hall who confirmed the diagnosis of JATD in all cases.

DNA samples from the proband (II.10) and both parents (I.1 and I.2) were collected by Dr Valerie Cormier-Daire in Paris and aliquots were subsequently sent to our lab.

(xi) Family 11

This consanguineous Turkish family (Fig. 2.14) was clinically assessed by Dr Beyhan Tuysuz. There were two affected siblings (II.1 and II.2). The parents (I.1 and I.2) were first cousins once removed. Both affected children had short stature, a narrow thorax and postaxial polydactyly. No major respiratory problems occurred in the neonatal period. Now aged of 14 years old (II.1) and 4 years old (II.2), no other complications have been reported.

Initially the two affected children were clinically and radiologically diagnosed as having JATD and were included in our study. DNA samples were obtained from the two affected sibs (II.1 and II.2) and both parents (I.1 and I.2). When copies of the radiographs were obtained, however, it became clear that the diagnosis was not straightforward. Radiographs were reviewed by Prof. Christine Hall, our collaborator Dr Valerie Cormier-Daire and Prof. Geert Mortier and they all agreed that the diagnosis was not typical JATD. Prof. Hall suggested acromesomelic dysplasia, Maroteaux type, while Dr Cormier-Daire suggested an atypical form of EVC.

(xii) Family 12

This distantly consanguineous English family (Fig. 2.14) was clinically assessed by Dr Peter Turnpenny, Consultant Clinical Geneticist at the Royal Devon and Exeter Hospital, Exeter. The parents (I.1 and I.2) were said to be almost certainly consanguineous because both the father's parents and the mother's grandparents originated from the same area at the tip of Cornwall and the same surname occurs on both sides of the family. There was one affected fetus (II.2) terminated at 19 weeks for probable JATD and two unaffected sibs (II.1 and II.3). Post-mortem findings were compatible both with JATD and with EVC. The fetus had short ribs and polydactyly. There was no evidence of congenital heart disease. Kidneys appeared normal and there were no gross abnormalities of the liver.

DNA samples were obtained from the fetus (II.2), the two unaffected children (II.1 and II.3) and both parents (I.1 and I.2).

2.1.1.2 Clinical details of non-consanguineous families

(xiii) Family 13

This Italian family (Fig. 2.16) was clinically assessed by Prof. Romano Tenconi, Consultant Clinical Geneticist at Padua Hospital, Italy. The proband (III.1) is the only child of unrelated parents. The two families have been living in the same area near Vicenza for over a century and the surname of the maternal grandmother, although a very common surname in the region, is the same as that of the father, so the two families may be distantly related.

Patient III.1 (Fig. 2.17A) was severely affected and had short stature, shortening of the upper and lower limbs and a very narrow chest. She was admitted to hospital several times in her first year of life due to dyspnea and feeding difficulties. Congenital hepatic fibrosis was diagnosed at 5 months of age and since then she has been treated with

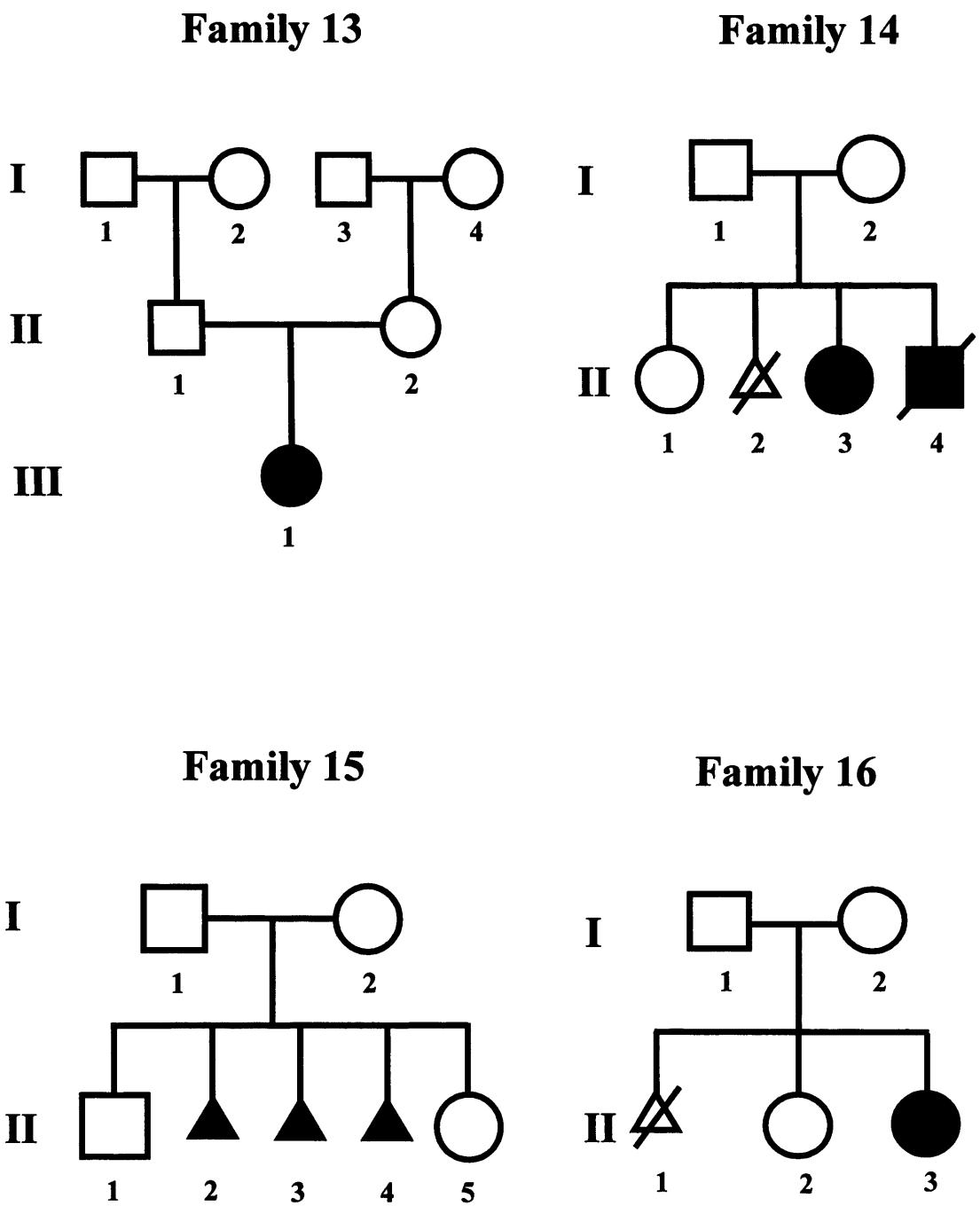


Figure 2.16 Pedigree drawing of Families 13, 14, 15 and 16, four non-consanguineous families with JATD. Solid symbols represent affected individuals.

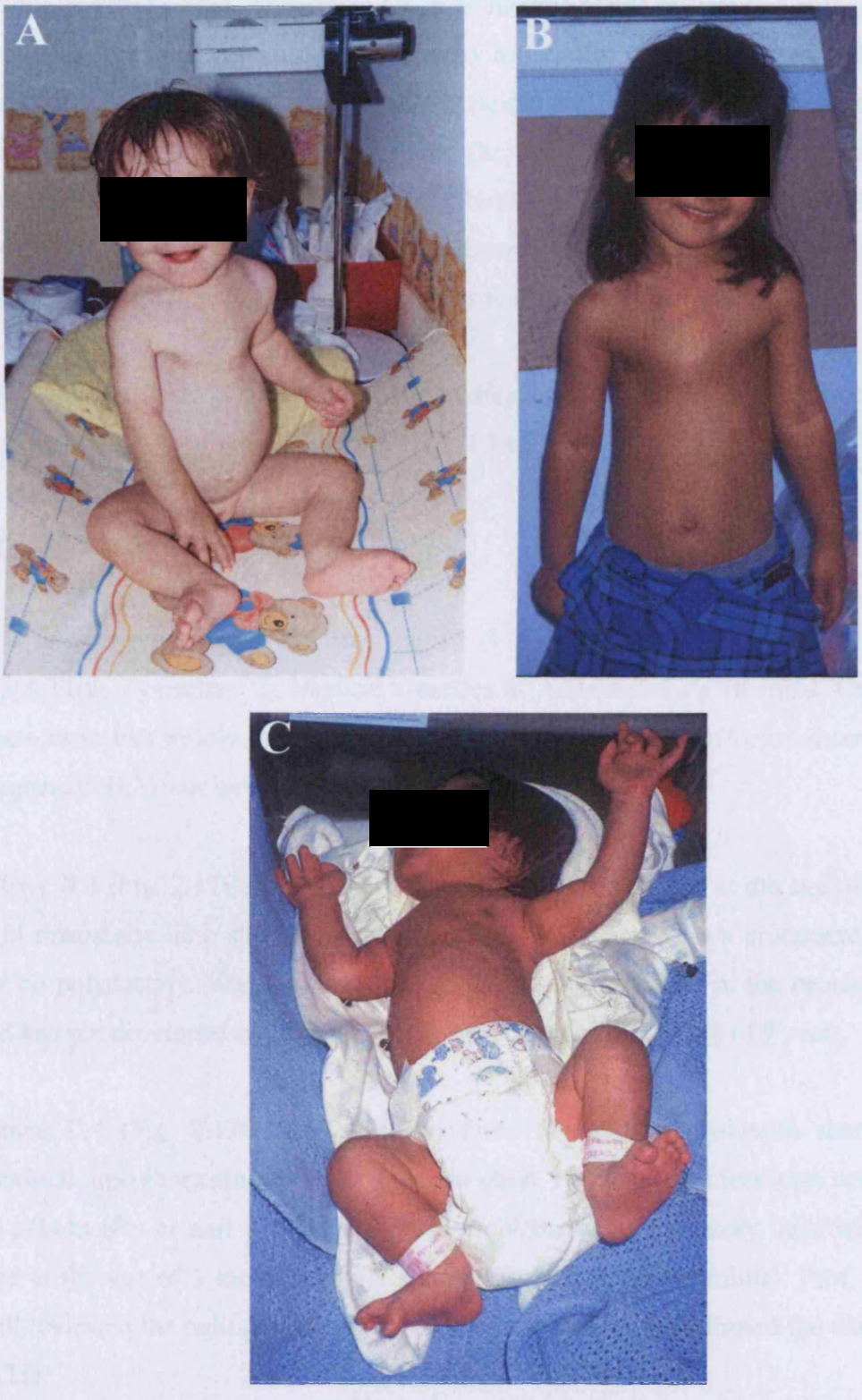


Figure 2.17 Clinical photographs of patients III.1 from Family 13 (A), II.3 (B) and II.4 (C) from Family 14 showing typical clinical findings of JATD. Limb shortening and a very narrow chest are present in all three cases.

Ursodesoxycholic acid. Abdominal US at 10 months of age revealed an enlargement of the spleen. Exocrine pancreatic insufficiency had appeared since the age of 14 months. Now aged 2½ years, she is <0.4th centile in height and 0.4th centile in weight and she is suffering from recurrent chest infections (bronchiolitis) and episodes of respiratory distress. The skeletal radiographs taken at birth supported the diagnosis of JATD. She had a narrow thorax with short thin ribs, rhizomelic limb shortening with no significant alterations in the metaphyses and a hypoplastic iliac crest.

DNA samples were obtained from the unaffected child (III.1), both parents (II.1 and II.2) and both sets of grandparents (I.1, I.2, I.3 and I.4).

(xiv) Family 14

This non-consanguineous Pakistani family (Fig. 2.16) was clinically assessed by Dr Helen Firth, Consultant in Medical Genetics at Addenbrooke's Hospital, Cambridge. There were two mildly affected sibs (II.3 and II.4) and one unaffected sister (II.1). A pregnancy (II.2) was terminated at 7-8 weeks of gestation.

Patient II.3 (Fig. 2.17B) had short stature (2nd centile for height at the age of 3 years), mild rhizomelic limb shortening and a narrow upper chest with a prominent sternum, but no polydactyly. She had no particular respiratory problems in the neonatal period and had not developed any extra-skeletal complications by the age of 9 years.

Patient II.4 (Fig. 2.17C and 2.18) was more severely affected with short stature, proximal limb shortening and a long narrow chest. His hands and feet were normal with no polydactyly or nail dysplasia. He had problems with respiratory insufficiency and died at the age of 3 months from respiratory infection (bronchiolitis). Prof. Christine Hall reviewed the radiographs of both affected children and confirmed the diagnosis of JATD.
DNA samples were first obtained from the two affected children (II.3 and II.4) and both parents (I.1 and I.2) and subsequently from the unaffected sib (II.1).

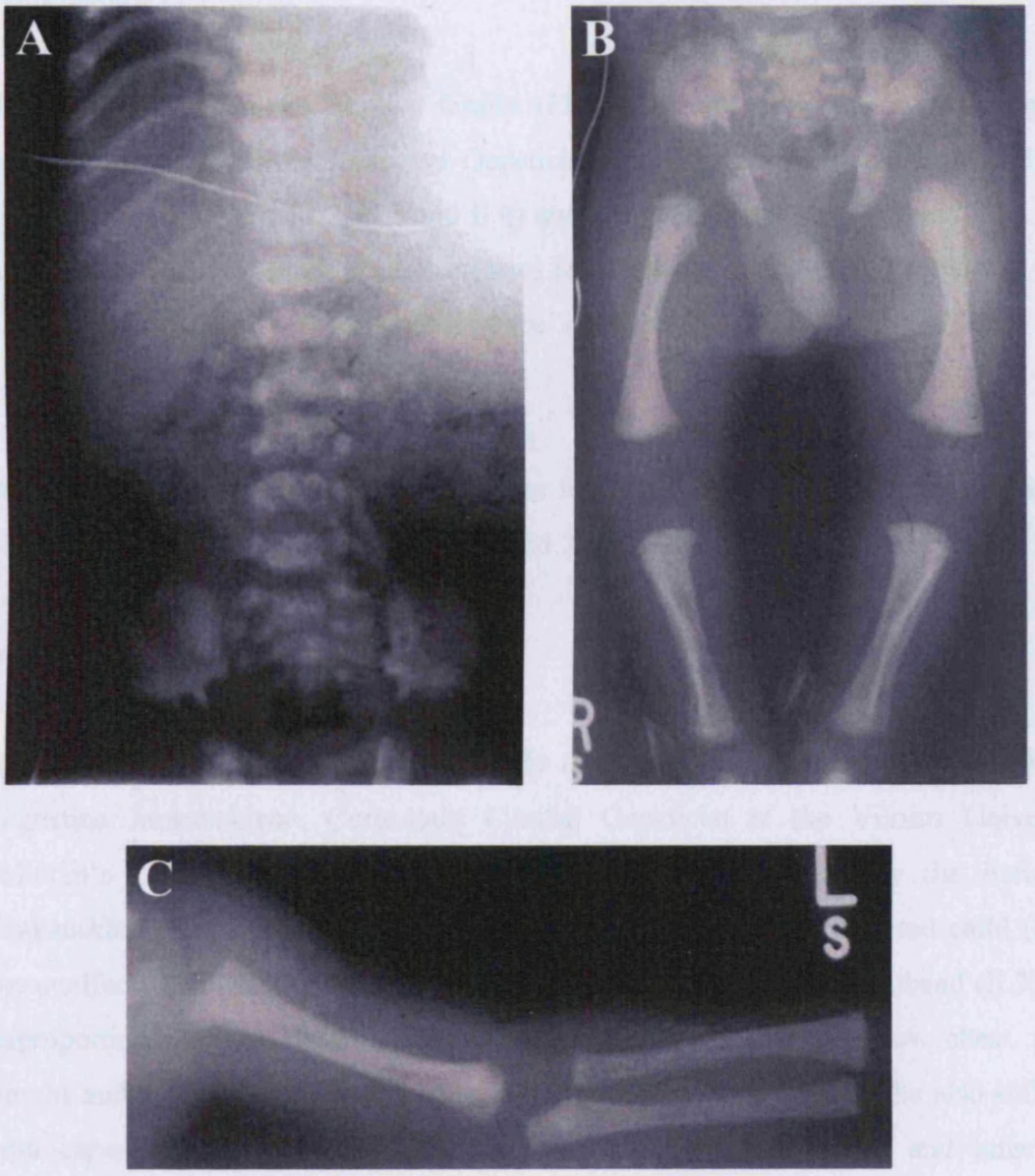


Figure 2.18 Radiographs of patient II.4 from Family 14 showing flaring of the ribs (A), trident appearance of the acetabulae (B) and rhizomelic limb shortening of the upper and lower limbs with flaring of the metaphyses (B and C).

(xv) Family 15

This non-consanguineous English family (Fig. 2.16) was clinically assessed by Dr Sarah Smithson, Consultant Clinical Geneticist at St Michael Hospital, Bristol. There were 3 affected fetuses (II.2, II.3 and II.4) and two unaffected sibs (II.1 and II.5). The three fetuses had post-mortem examinations and fetal radiographs were consistent with a diagnosis of JATD. Radiographs were reviewed by Prof. Christine Hall who confirmed the diagnosis.

DNA samples were obtained from the three fetuses (II.2, II.3 and II.4), the unaffected sibs (II.1 and II.5) and both parents (I.1 and I.2).

(xvi) Family 16

This non-consanguineous Lithuanian family (Fig. 2.16) was clinically assessed by Dr Augustina Jankauskiené, Consultant Clinical Geneticist at the Vilnius University Children's Hospital, Vilnius, Lithuania and has been reported in the literature (Jankauskiene and Bernatoniene, 2000). There was one severely affected child (II.3), one unaffected sib (II.2) and a termination of pregnancy (II.1). The proband (II.3) had disproportionate short stature, short limbs, a short neck and a narrow chest. At 4 months and at 2 years of age she had episodes of respiratory failure. She also suffered from ear-nose-throat infections and experienced severe pneumonia and infectious mononucleosis. By 7 months, she developed renal insufficiency. Her liver was enlarged, but liver function tests were normal. Eye examination was normal. Skeletal radiographs showed a long, narrow thorax with short horizontal ribs, short limbs and distal phalanges and trident appearance of the acetabulae in the pelvis.

DNA samples were collected by Dr Colin Johnson in Birmingham and aliquots from all four members of the family were subsequently sent to our lab.

(xvii) Family 17

This non-consanguineous Lebanese family (Fig. 2.19) was clinically assessed by Dr Imad Kaddoura, Chief Plastic Surgeon at the American University of Beirut Medical Center, Beirut, Lebanon. The proband (II.1) has been reported in the literature (Kaddoura *et al.*, 2001). He was the third child of unaffected parents (I.1 and I.2). There were two unaffected sisters (II.2 and II.3). The proband was severely affected with a small chest and short long bones in the lower limbs and died at 18 months of age despite thoracoplasty. There was no renal or hepatic involvement.

Radiographs were reviewed by Prof. Christine Hall who confirmed the diagnosis of JATD. He had a very narrow thorax with short ribs. In the pelvis, the acetabular roofs showed a trident configuration with medial and lateral spurs. The long bones in the lower limbs were short with irregular metaphyses and premature ossification of the capital femoral epiphyses.

DNA samples were obtained from the two unaffected sibs (II.2 and II.3) and both parents (I.1 and I.2).

(xviii) Family 18

This non-consanguineous Scottish family (Fig. 2.19) was clinically assessed by Dr John Dean, Consultant Clinical Geneticist at the Grampian University Hospital, Aberdeen, Scotland. There was one unaffected child (II.2) and two affected children (II.1 and II.3). The first (II.1) was a severely affected girl who died aged 6 hours (II.1). She was small and had a bell shaped chest and pulmonary hypoplasia. The second (II.3) was a less severely affected girl. She had mild skeletal involvement with short stature, rhizomelic limb shortening and a long narrow bell-shaped chest. She had no neonatal respiratory problems and no renal or hepatic involvement. Radiographs of the affected child (II.3) were reviewed at the London Dysmorphology Club and the diagnosis of JATD was confirmed. She had slightly bowed femurs and short humeri.

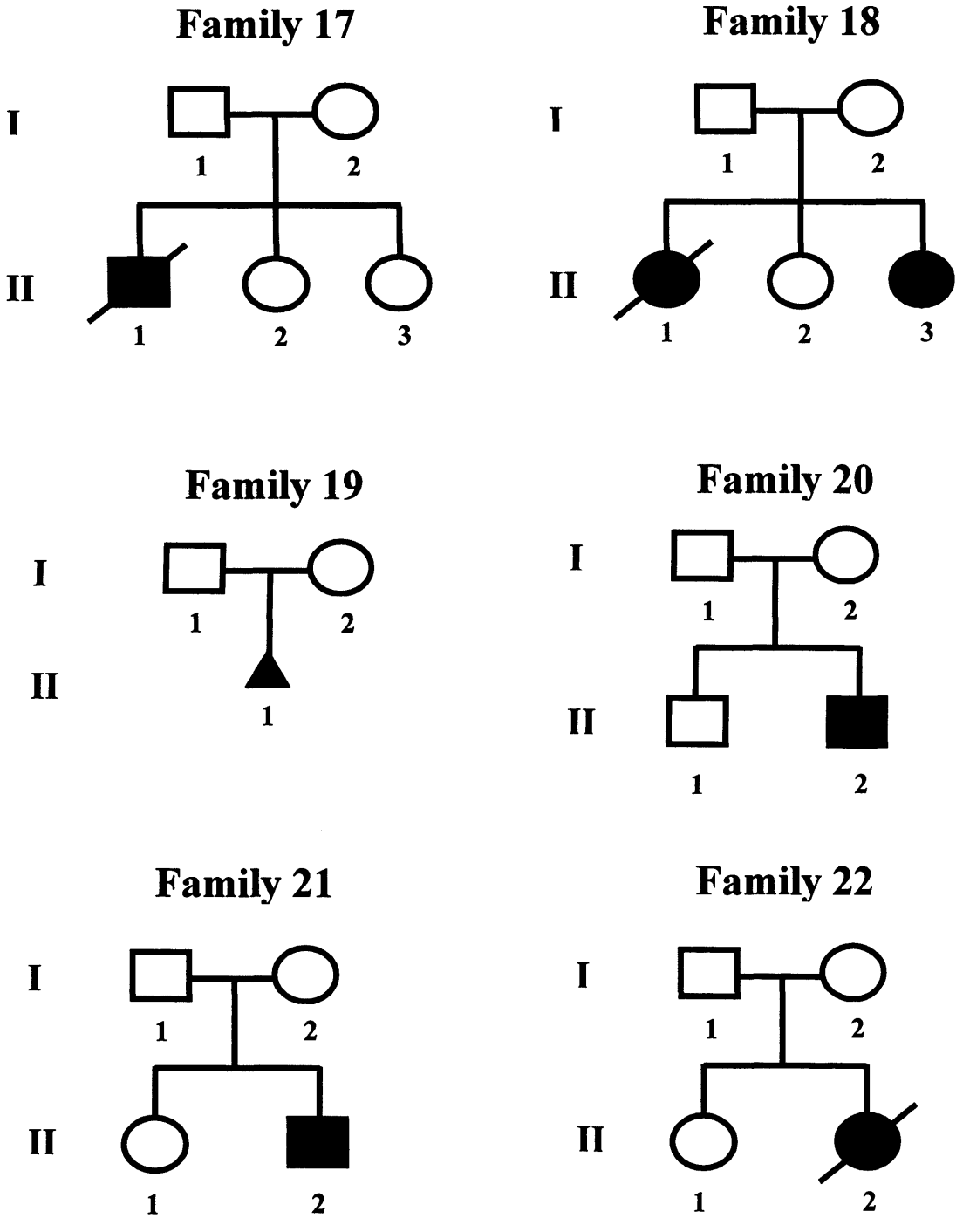


Figure 2.19. Pedigree drawing of Families 17-22, six non-consanguineous families with JATD. Solid symbols represent affected individuals.

DNA samples were obtained from the parents (I.1 and I.2) and one affected child (II.3).

(xix) Family 19

This non-consanguineous Belgian family (Fig. 2.19) was clinically assessed by Dr Jean-Pierre Fryns, Consultant Clinical Geneticist at the Center for Human Genetics, Leuven, Belgium. There was one affected female fetus (II.1). Post-mortem examination was compatible with JATD. She had a narrow thorax and short limbs, but no polydactyly or dysmorphic features. Radiographs showed a narrow thorax with short ribs, shortening of the long bones of the limbs and irregular metaphyses with spikes, but no bowing of the femora.

DNA samples were obtained from both parents (I.1 and I.2) and from the fetus (II.1).

(xx) Family 20

This non-consanguineous Portuguese family (Fig. 2.19) was clinically assessed by Dr Ana Medeira, Consultant Clinical Geneticist at the Santa Maria Hospital, Lisbon, Portugal. The proband (II.2) was the second child of healthy parents (I.1 and I.2). Short limbs were noted on antenatal US. He had short stature, mild rhizomelic limb shortening and a narrow chest. He had episodes of respiratory distress (asthma) in the first two years of life. There was no renal or hepatic involvement. Radiological examination showed short ribs with anterior flaring, small ilia with square wings, trident acetabulae and irregular metaphyses of the long bones. The films were reviewed by Prof. Christine Hall who confirmed the diagnosis of JATD.

DNA samples were obtained from the proband (II.2), the unaffected sib (II.1) and both parents (I.1 and I.2).

(xxi) Family 21

This non-consanguineous Australian family (Fig. 2.19) was clinically assessed by Dr Ravi Savarirayan, Consultant Clinical Geneticist at the Royal Children's Hospital Genetics Clinic, Victoria, Australia. The proband (II.2), the second child of healthy unrelated parents (I.1 and I.2), was mildly affected. He had short stature, a narrow chest and soft tissue syndactyly of the fingers and toes. He had no respiratory problems and normal renal function, but had abnormal liver function tests and cholestasis and was treated with bile acids.

DNA samples were obtained from the proband (II.2), the unaffected sib (II.1) and both parents (I.1 and I.2).

(xxii) Family 22

This non-consanguineous Italian family (Fig. 2.19) was clinically assessed by Dr Carlo Poggiani, Consultant Clinical Geneticist at Cremona Hospital, Italy and has been reported in the literature (Poggiani *et al.*, 2000). The proband (II.2) had severe skeletal involvement and died 24 hours after birth of respiratory failure. She was small and had short limbs and a very narrow chest. Cranial and renal US were normal. Post-mortem examination and fetal radiographs confirmed the diagnosis of JATD. She had a narrow thorax with short ribs and a trident appearance of the acetabular roofs in the pelvis.

DNA samples were obtained from the unaffected sib (II.1) and both parents (I.1 and I.2).

(xxiii) Family 23

This non-consanguineous English family (Fig. 2.20) was clinically assessed by Dr Tessa Homfray, Consultant Clinical Geneticist at St George's Hospital, London. There was one affected fetus (II.1) terminated at 22 weeks of gestation. Diagnosis of JATD

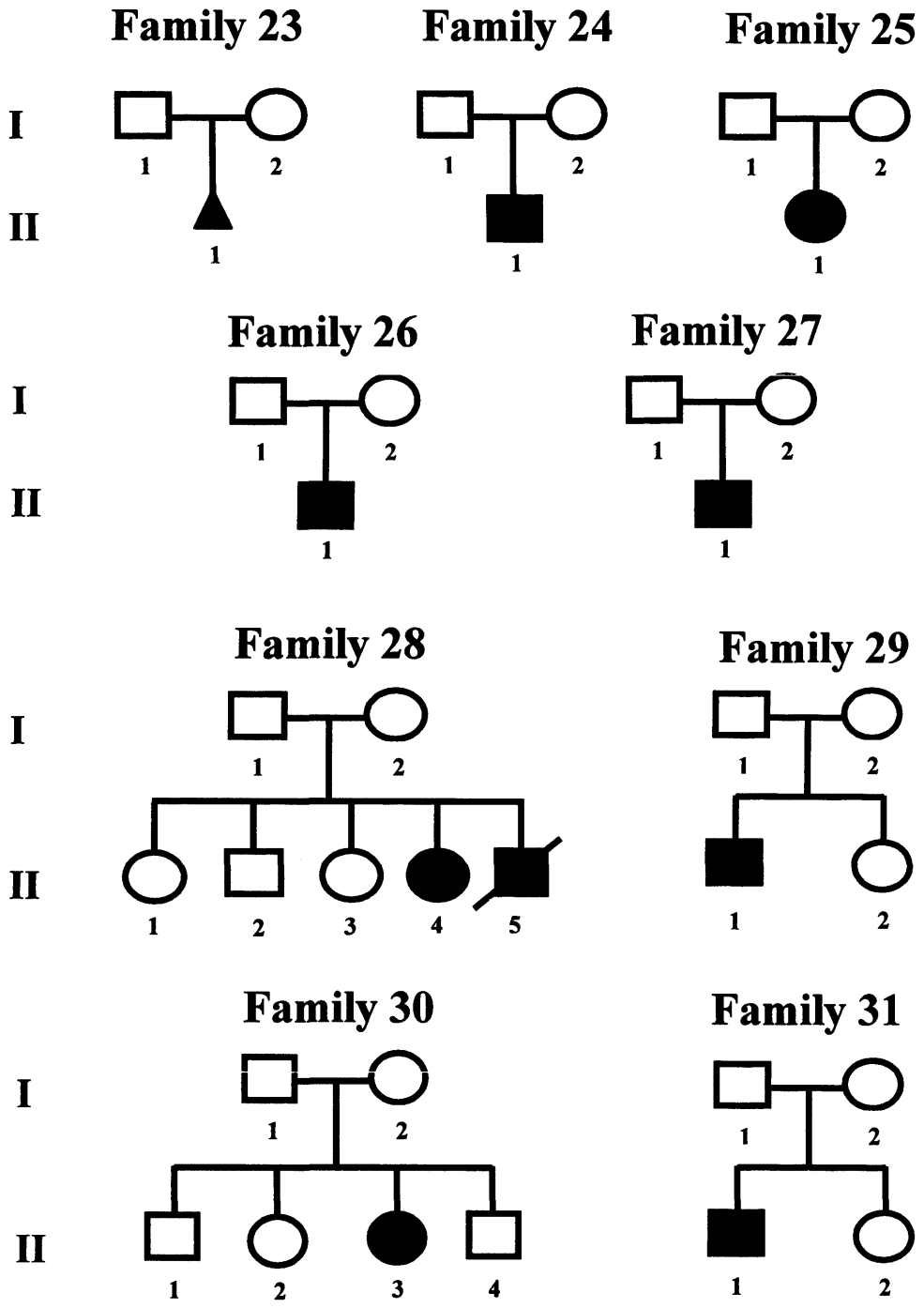


Figure 2.20 Pedigree drawing of Families 23-31, nine non-consanguineous families with JATD. Solid symbols represent affected individuals.

was confirmed at post-mortem examination and on fetal radiographs reviewed by Prof. Christine Hall. DNA samples were obtained from the parents (I.1 and I.2).

(xxiv) Family 24

This non-consanguineous family from Kosovo (Fig. 2.20) was clinically assessed by Dr Radovan Bogdanovic, Consultant Clinical Geneticist at the Institute of Mother and Child Health, Belgrade, Serbia. There was only one affected child (II.1) with skeletal, renal and hepatic involvement. DNA samples were obtained from the proband and his parents (I.1 and I.2).

(xxv) Family 25

This non-consanguineous family from Serbia (Fig. 2.20) was also clinically assessed by Dr Radovan Bogdanovic. There was only one affected child (II.1) with skeletal, renal, retinal and hepatic involvement. DNA samples were obtained from the proband and his parents (I.1 and I.2).

(xxvi) Family 26

This non-consanguineous family from Serbia (Fig. 2.20) was also clinically assessed by Dr Radovan Bogdanovic, Belgrade, Serbia. There was only one affected child (II.1). No clinical information about the proband has been obtained yet. DNA samples were obtained from the proband and his parents (I.1 and I.2).

(xxvii) Family 27

This non-consanguineous family from Serbia (Fig. 2.20) was also clinically assessed by Dr Radovan Bogdanovic, Belgrade, Serbia. There was only one affected child (II.1). No clinical information about the proband has been obtained yet. DNA samples were obtained from the proband and his parents (I.1 and I.2).

(xxviii) Family 28

This non-consanguineous Dutch family (Fig. 2.20) was clinically assessed by Dr Ben Hamel and Prof. Han Brunner, Consultant Clinical Geneticists at the Department of Human Genetics, Nijmegen, Netherlands. There were two affected children (II.4 and II.5) and three unaffected sibs (II.1, II.2 and II.3). Patient II.5 died 38 days after birth. DNA samples were obtained from patient II.4, the three unaffected children and their parents (I.1 and I.2).

(xxix) Family 29

This non-consanguineous Dutch family (Fig. 2.20) was also clinically assessed by Dr Ben Hamel and Prof. Han Brunner. There was one affected child (II.1) and one unaffected sib (II.2). DNA samples were obtained from the proband (II.1), the unaffected sib (II.2) and both parents (I.1 and I.2).

(xxx) Family 30

This non-consanguineous Dutch family (Fig. 2.20) was also clinically assessed by Dr Ben Hamel and Prof. Han Brunner. There was one affected child (II.3) and three unaffected sibs (II.1, II.2 and II.4). DNA samples were obtained from the proband (II.3), the three unaffected children and their parents (I.1 and I.2).

(xxxi) Family 31

This non-consanguineous Dutch family (Fig. 2.20) was also clinically assessed by Dr Ben Hamel and Prof. Han Brunner. There was one affected child (II.1) and one unaffected sib (II.2). DNA samples were obtained from the proband (II.1), the unaffected sib (II.2) and both parents (I.1 and I.2).

2.1.2 CLS patient samples

The project started when an affected child from a consanguineous family (Family A) was seen by clinicians working in the Clinical and Molecular Genetics Unit at Great Ormond Street Hospital. DNA samples were later obtained from three further consanguineous families (B, C, D) and three non-consanguineous families (E, F, G). DNA samples from whole Family B were obtained through Dr Nursel Elcioglu, Consultant Clinical Geneticist at Marmara University Hospital, Istanbul, Turkey. DNA samples from Family C and Family D were obtained through Prof. Samia A. Temtamy, Professor of Human Genetics at the National Research Centre of Cairo, Egypt.

A DNA sample from an affected child from a non-consanguineous family (Family E) was obtained through Prof. Jean-Pierre Fryns, Consultant Clinical Geneticist at the University of Leuven, Belgium. Samples from two further non-consanguineous families (Family F and Family G) were obtained later through Dr Nursel Elcioglu.

2.1.2.1 Clinical details of CLS families

(a) Family A

This consanguineous Asian family (Fig. 2.21) was clinically assessed by Prof. Robin Winter, Consultant Clinical Geneticist at Great Ormond Street Hospital. The proband (II.2) was the second child of first-cousin parents. Her parents (I.1 and I.2) and older sister (II.1) were unaffected. She had severe symmetrical involvement of both hands, with total soft tissue syndactyly and shortening of all fingers (Fig. 2.22A and B), as well as bilateral radio-ulnar synostosis, preventing pronation and supination of the forearms. Radiographs of her upper limb (Fig. 2.22C and D) were reviewed by Prof. Christine Hall, Consultant Radiologist in the Department of Radiology at Great Ormond Street Hospital and were consistent with a diagnosis of CLS. In the feet, the only abnormality was partial soft tissue syndactyly between the second and third toes.

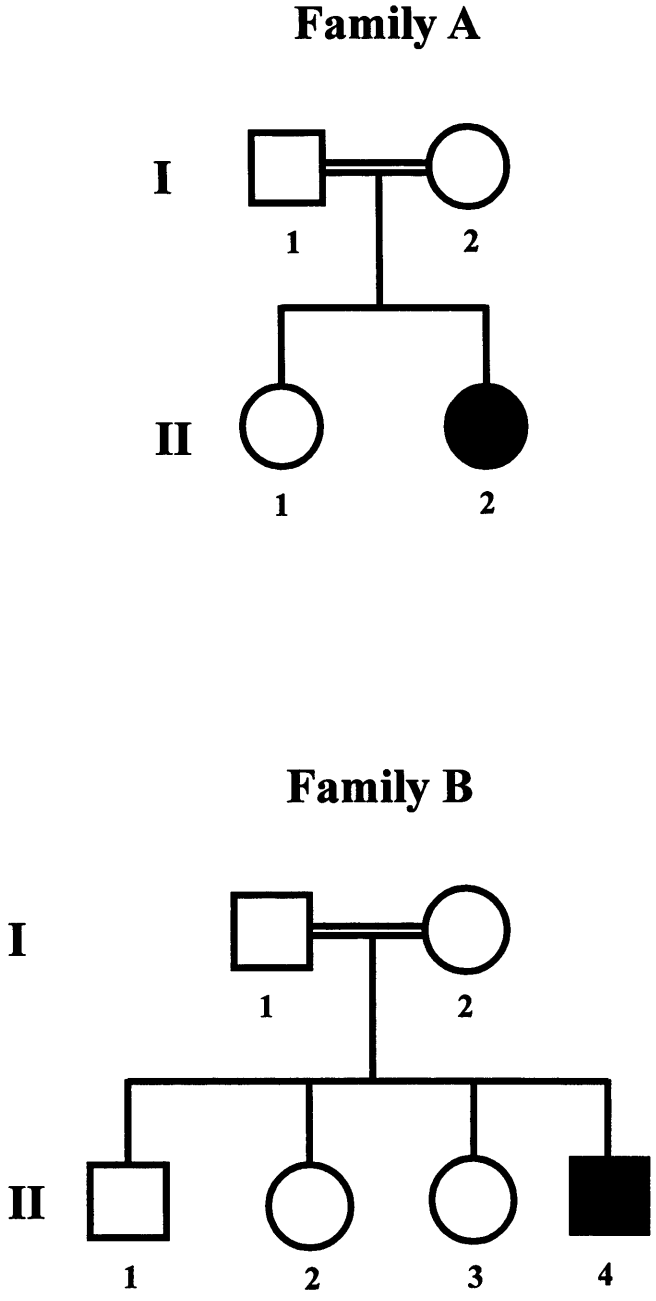


Figure 2.21 Pedigree drawing of Families A and B, two consanguineous families with Cenani-Lenz syndrome. Solid symbols represent affected individuals.

Figure 2.22 Clinical photographs and radiographs of the hands and arms of patient II.2 from Family A. The rudimentary thumbs, index and little fingers have separate nails, but the middle and ring fingers shared a single fused nail (A, B). The phalanges and metacarpals of both hands are malformed and fused, and there is bilateral fusion of the entire shafts of the radius and ulna, with some dislocation of the elbow joints (C, D).

Interestingly an abdominal US showed bilateral renal hypoplasia, more severe on the left then on the right. A short report describing this child has already been published (Bacchelli *et al.*, 2001), and a copy is included at the back of this thesis. She is now under the care of the Plastic Surgery team at Great Ormond Street Hospital, and is undergoing a series of operations on her hands. DNA samples were obtained from both parents (I.1 and I.2) and the affected child (II.2).

(b) Family B

This consanguineous Turkish family (Fig. 2.21) was clinically assessed by Dr Nursel Elcioglu, Consultant in Clinical Genetics at Marmara University Hospital, Istanbul, Turkey, and has already been reported in the literature (Elcioglu *et al.*, 1997). The proband (II.4) was the fourth child of healthy first-cousin parents (I.1 and I.2). His three older sibs (II.1, II.2 and II.3) were unaffected. He had bilateral symmetrical abnormalities of the hands with four short, partly syndactylous fingers with nails (Fig. 2.23A and B). The radii and ulnae were slightly short, but there was no radio-ulnar synostosis. The only abnormality in the feet was bilateral syndactyly between the second and third toes (Fig. 2.23A). No other malformations apart from the limbs were detected and his developmental progress was normal. DNA samples were obtained from both parents (I.1 and I.2), the affected child (II.4) and the three unaffected sibs (II.1, II.2 and II.3).

(c) Family C

This consanguineous Egyptian family (Fig. 2.24) was clinically assessed by Prof. Samia A. Temtamy, Professor of Human Genetics at the National Research Centre of Cairo, Egypt, and has also been reported in the literature (Temtamy *et al.*, 2003). The proband (III.4) was the fourth child of double first cousin parents. She had bilateral symmetrical malformations of both upper limbs and feet and facial dysmorphism. She had short forearms, disorganised fingers with two malformed central digits, restricted supination and pronation at the elbows and restricted flexion and extension at the wrist.

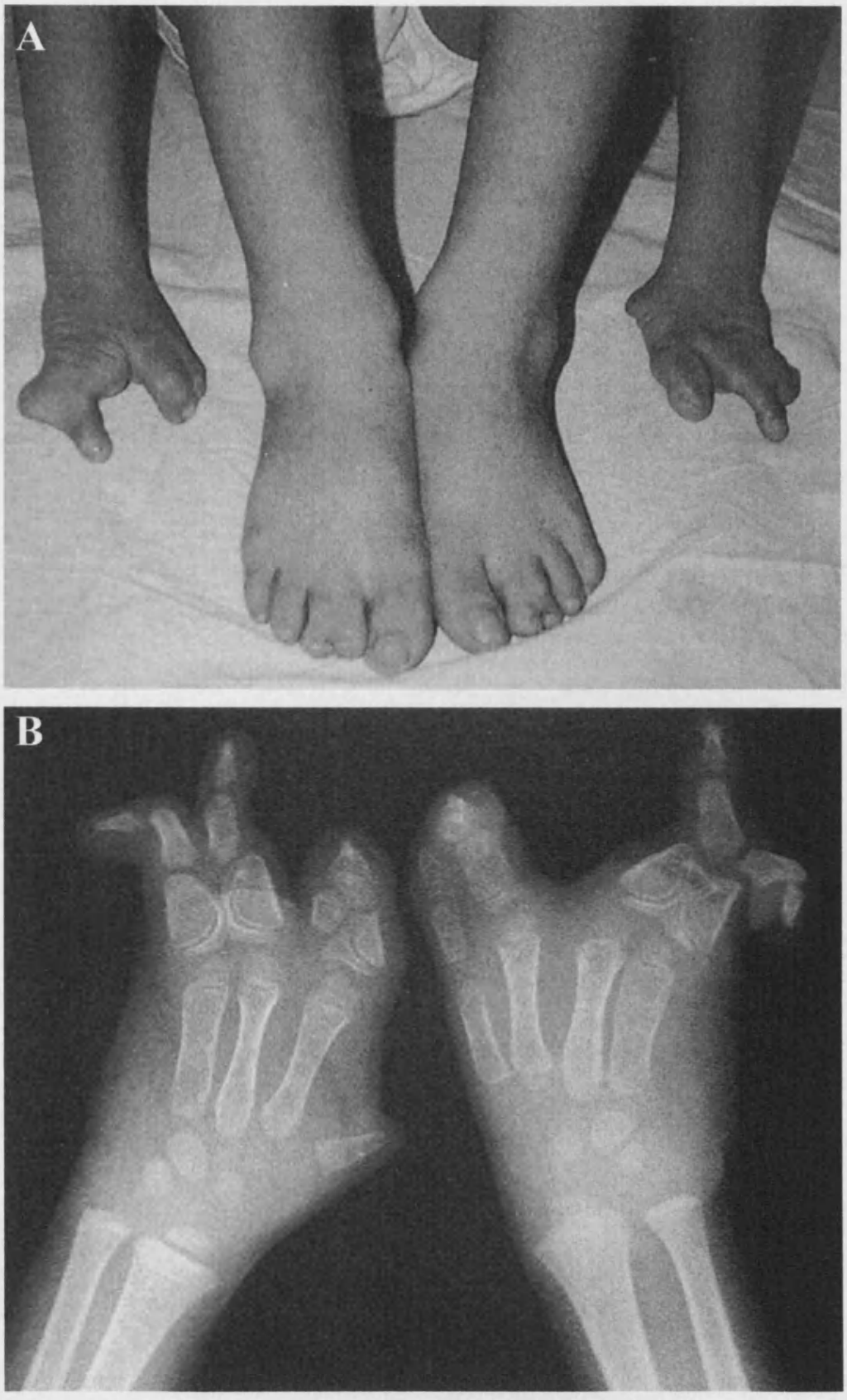


Figure 2.23 Clinical photograph and radiograph of the limbs of patient II.4 from Family B. In the hands, there is bilateral soft tissue syndactyly of all the digits, with shortening, fusion or absence of many of the phalanges and metacarpals (A and B). In the feet, there is bilateral syndactyly between the second and third toes.

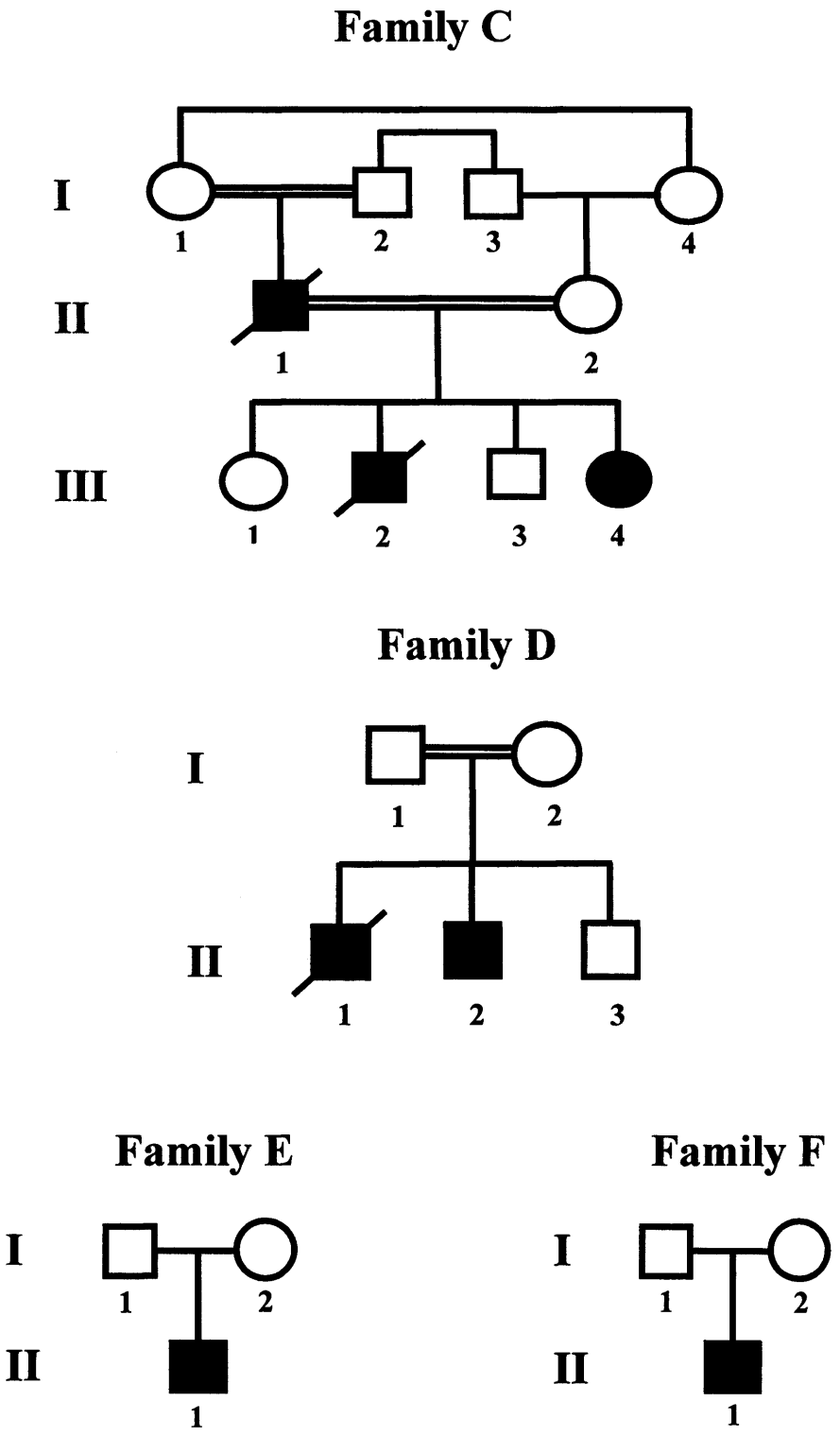


Figure 2.24 Pedigree drawing of Families C, D, E and F, two consanguineous and two non-consanguineous families with Cenani-Lenz syndrome. Solid symbols represent affected individuals.

Radiological examination revealed fusion of the radius and ulna, fusion of the carpals and of the first and second metacarpals and of the phalanges. On each foot only four toes were visible with a wide gap between the first and second toes. X-rays showed bilateral asymmetrical malformations with fusion of some metatarsals and proximal phalanges, absence of middle phalanges and biphalangeal second to fifth toes. Facial dysmorphism consisted of a broad, prominent forehead, hypertelorism, downslanting palpebral fissures, a depressed nasal bridge, a short nose, prominent central incisors and a bow shaped upper lip. Tooth abnormalities like enamel hypoplasia and premature loss of permanent teeth were described. She also had a high arched palate and partial inferior ankyloglossia. Both kidneys were of normal size on abdominal US.

The father (II.1) had died and was said to have upper and lower limb malformations and facial dysmorphism similar to the proband's. His parents were unaffected first cousins suggesting a quasi-dominant inheritance in this family. The proband had two unaffected sibs (III.1 and III.3) and a brother (III.2) who was also said to be similarly affected and died at the age of 5 months from measles. DNA samples were obtained from the proband (III.4) and the proband's mother (II.2).

(d) Family D

This consanguineous Egyptian family (Fig. 2.24) was clinically assessed by Prof. Samia A. Temtamy and has been reported in the literature together with Family C (Temtamy *et al.*, 2003). The proband (II.2) was the second child of first cousin unaffected parents. He had bilateral symmetrical malformations of upper limbs, feet and facial dysmorphism. Both upper limbs were severely affected with bilateral symmetrical shortening of the forearms, limited extension, pronation and supination at the elbows and disorganised fingers. Radiological examination showed partial radio-ulnar synostosis and fusion of proximal phalanges into two cone shaped structures. Only four metacarpals were identified. The feet were small with only four vestigial toes bilaterally, syndactyly of the first to third toes, a gap between the third and last toe, and hypoplastic nails. X-rays showed absence of one metatarsal and some phalanges. He

had facial features similar to those of proband III.4 from Family C including hypertelorism and orodental abnormalities such as partial tongue-tie, high arched palate, and delayed eruption of teeth. Both kidneys were normal on abdominal US.
The sibling (II.3) was unaffected but a baby (II.1) who was said to have severe limb abnormalities, cleft lip and palate and ambiguous genitalia, died a few hours after birth. DNA samples were obtained from four members of the family (I.1, I.2, II.2, II.3).

Family C and D had atypical features such as facial dysmorphism, dental abnormalities and partial tongue-tie, which have never been described before in CLS patients.

(e) Family E

This non-consanguineous Belgian family (Fig. 2.24) was clinically assessed by Prof. Jean-Pierre Fryns, Consultant Clinical Geneticist at the University of Leuven, Belgium and has also been reported (De Smet *et al.*, 1992). The proband (II.1) was the only child of unaffected unrelated parents. His hands were hypoplastic with complete syndactyly between all fingers, fusion of some of the nails, and shortening, fusion or absence of many of the phalanges and metacarpals (Fig. 2.25A, B and C). A series of surgical operations were performed to release the syndactyly. The forearms were normal, except for a dislocation of the radial head of the right elbow, and the feet were normal too. He had no other abnormalities apart from the hand malformations. A DNA sample was obtained from the proband (II.1).

(f) Family F

This non-consanguineous Turkish family (Fig. 2.24) was clinically assessed by Dr Nursel Elcioglu and has been reported in the literature (Seven *et al.*, 2000). The proband (II.1) was the first child of healthy parents. She had upper and lower limb abnormalities typical of CLS. She had mesomelic shortening of both arms, symmetrical complete syndactyly of all the fingers and of the second to fifth toes on the right foot and of the second and third toes on the left foot. Radiological examination showed

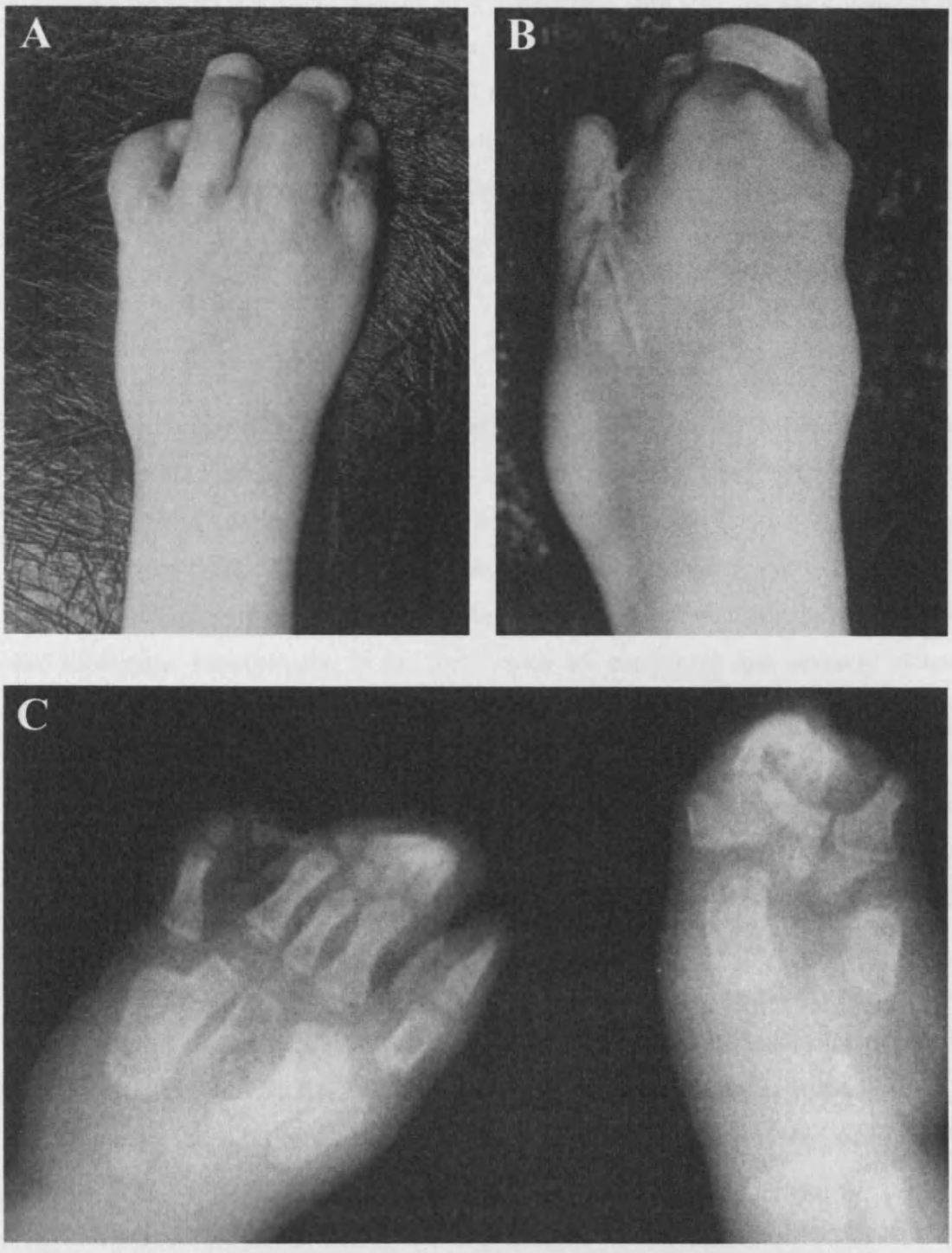


Figure 2.25 Clinical photographs and radiograph of the hands of patient II.1 from Family E. There is soft tissue syndactyly of the fingers and fusion of some of the nails, more severe on the right (B) than on the left (A). Many of the phalanges and metacarpals are malformed, fused or missing, especially on the right, where the phalanges form a disorganized mass of bone.

bilateral radio-ulnar synostosis, disorganisation of metacarpals and phalanges and synostosis between the fourth and the fifth metatarsals. She also had extra features not previously reported in CLS patients such as vertebral abnormalities (scoliosis of the thoracic spine and hemivertebrae in T10, T11 and T12); rib anomalies (fork-like 5^{th} rib, wide separation between the 6^{th} and 7^{th} rib and thickened 8^{th} rib) and mixed hearing loss. DNA samples were obtained from the proband (II.1) and from both parents (I.1 and I.2).

(g) Family G

This Turkish family (Fig. 2.26) was also clinically assessed by Dr Nursel Elcioglu and has been reported in the literature (Percin and Percin, 2003). The proband (IV.11) was the tenth child of unaffected, unrelated parents. There were eleven unaffected siblings. She had severe typical CLS features in her upper limbs including syndactyly between all digits and fusion of the entire shaft of the right radius and ulna preventing pronation and supination. Interestingly, in the feet, which are commonly less severely affected than the hands in CLS, she had an unusual form of duplication of the phalanges of both first toes and of the right second toe, a feature which has not previously been reported in CLS. She had hypoplasia of the middle and distal phalanges of the second to fifth toes and the same feature was present in her mother (III.2). The proband also had congenital cataract in the right eye, again a feature not previously associated with CLS. A half-cousin of the proband (V.2) whose parents were related, had an unusual form of mesoaxial syndactyly characterised by bilateral synostosis of the phalanges of the third and fourth fingers and partial soft tissue syndactyly between the right second and third toes, which does not fit any recognised type of syndactyly. This is very similar to a sporadic case, the child of unrelated parents, reported by Temtamy and McKusick, who classed it as either a severe form of syndactyly type 1 (SD1) or a 'new' type of syndactyly (Temtamy and McKusick, 1978). Percin *et al.* (1998) have described a similar severe mesoaxial syndactyly in three patients from a large inbred pedigree with SD1, the children of affected first cousin parents. Their severe phenotype was thought to represent a homozygous form of SD1 (Percin *et al.*, 1998).

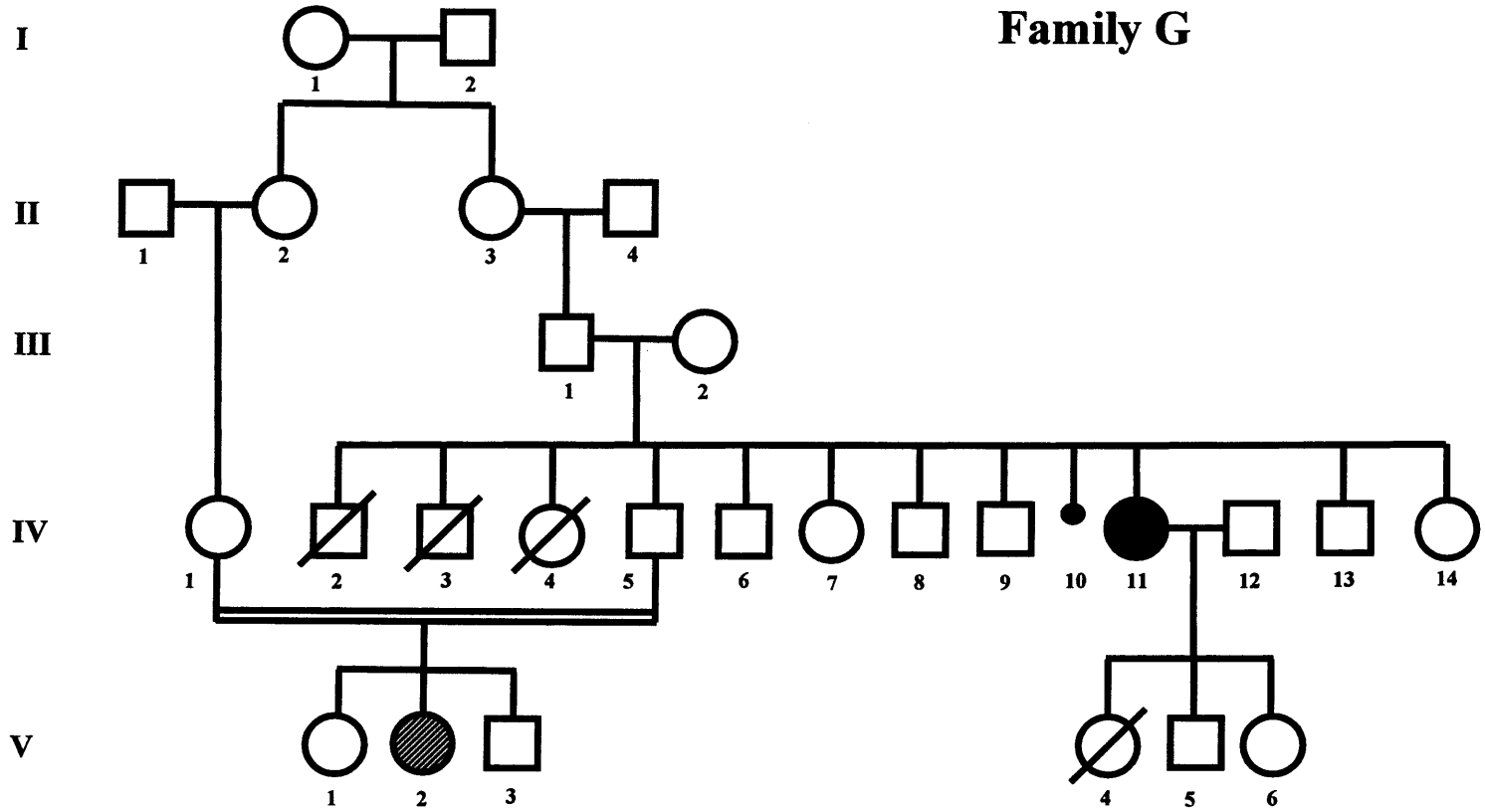


Figure 2.26 Pedigree drawing of Family G, a non-consanguineous family with Cenani-Lenz syndrome. Solid symbol represents affected individual. Striped symbol represents an unusual type of syndactyly.

DNA was obtained from the proband (IV.11), her parents (III.1, III.2), an unaffected sibling (IV.7) and the half-cousin (V.2).

2.2 Materials

2.2.1 Reagents

All reagents were of AnalaR grade, or the best grade obtainable, and were obtained from British Drug Houses (BDH) Laboratory Supplies or Sigma Aldrich, unless indicated otherwise. Glassware, solutions and media were autoclaved at 105kPa (15psi), 121°C for 20 minutes where necessary. Water was purified using a Milli-Ro 15 Water Purification System (Millipore SA), further purified where necessary using a Milli-Q Reagent Grade Water Ultrafiltration System (Millipore), and sterilised by autoclaving.

NuSieve GTG low melting point agarose was from FMC Bioproducts. Electrophoresis grade agarose and TEMED (N,N,N',N'-tetramethylethylenediamine) were obtained from Gibco-BRL Life technologies. Acrylamide stock solution (40% w/v, 29:1 acrylamide:N,N'-methylene bisacrylamide) and urea were from Bio-Rad Laboratories. Alconox detergent was from Alconox, Inc. Absolute ethanol and isopropanol were from Hayman Ltd. Bacto-agar, bacto-tryptone and yeast extract were from Difco Laboratories. Deionised formamide was from Eastman Kodak Co. RPMI 1640 medium, penicillin and streptomycin were from Gibco-BRL Life Technologies. Fetal bovine serum (FBS) and L-glutamine were from Imperial Laboratories Ltd. Amberlite resin and blue dextran/EDTA (50mg/ml and 25mM respectively) were from Perkin Elmer Applied Biosystems Division. Ultrapure dNTPs were supplied by Amersham Biosciences. Proteinase K was obtained from QIAGEN Ltd. NaOH pellets were from Fisons Scientific Equipment. Pellet Paint® NF Co-Precipitant was supplied by Novagen. 10x MegaBACE™ LPA Buffer was from Amersham Biosciences.

2.2.2 Non-standard materials

In addition to standard laboratory materials, use was made of the following special items. Falcon 15ml tubes, 50ml tubes, 2059 polypropylene tubes and tissue culture flasks were from Becton Dickinson Labware. Culture plates were from Bibby Sterilin Ltd. ART aerosol resistant barrier tips were from CLP. Sterile loops, cryotubes and filter sterilisation units were from Nalge Nunc International. 0.2ml individual PCR tubes and 0.2ml strip tubes were from ABgene. Polycarbonate 96 well microplates OMNIPLATE 96 ThermoHybaid were from Hybaid Ltd or ABgene (Thermofast ® 96 skirted plates). Plate seals HB-TD-TAPE Micro sealing sheets were from Hybaid Ltd. PCR Miniflex 0.2mm flat tips were from Sorenson BioScience, Inc. Glassware was obtained from Pyrex ® and Schott Duran. Plasticware was from Becton Dickson Lab ware or Bibby Sterilin Ltd. Micropore 0.2μm acrodisc ® syringe filters were supplied by Paul Gelman Labs. Plastic and paper sharks tooth combs were supplied by Perkin Elmer Applied Biosystems Division.

DNA electrophoresed on agarose gels was visualised by ethidium bromide fluorescence when the gels were illuminated at 300nm on a Chromato-Vue transilluminator (UVP, Inc.). Gels were photographed on a videocamera using Polaroid film (Polaroid Corporation) printing onto thermal paper (Mitsubishi Electric Corporation). PCR amplification was carried out using a T3 Thermocycler, a Personal Cycler™ (both from Biometra) or an Eppendorf Mastercycler gradient thermocycler (Eppendorf). DNA sequence analysis and microsatellite repeat analysis were performed on an ABI PRISM™ 377 DNA Sequencer (Perkin Elmer Applied Biosystems Division) or a MegaBACE 1000 DNA Sequencer (Amersham Biosciences).

2.2.2.1 Nucleotide size markers

The 50bp and 1kb DNA ladders were from Gibco-BRL Life Technologies. The GENESCAN-500™ TAMRA internal lane standard was from Perkin Elmer Applied

Biosystems Division. The MegaBACE™ ET400-R and MegaBACE™ ET550-R size standard were from Amersham Biosciences.

2.2.2.2 Oligonucleotides

Oligonucleotide primers were synthesized to order by MWG-Biotech, Genosys Biotechnologies Inc or QIAGEN Ltd. Fluorescent dye labelled oligonucleotides used in the genome-wide screen were from the Single Chromosome Scan linkage set (Research Genetics, Inc.), the Genome-wide Human Screening Set Version 8a (Research Genetics, Inc.) and the ABI PRISM™ Linkage Mapping Set (Perkin Elmer Applied Biosystems Division), available in the laboratory.

2.2.2.3 Enzymes

HpaII restriction endonuclease was from Promega Corporation. Restriction enzyme *AciI* and *E.coli* exonuclease III were from New England Biolabs Inc. AmpliTaq Gold™ was obtained from Perkin Elmer Applied Biosystems Division. HotStarTaq DNA Polymerase was from QIAGEN Ltd. Cloned *Pfu* DNA polymerase and Herculase™ Hotstart DNA polymerase were from Stratagene. Biopro™ *Taq* polymerase was from Bioline. All enzymes were used with the appropriate 10x buffers supplied by the manufacturers. ExoSAP-IT™ was from USB Corporation.

2.2.2.4 Commercial kits

ABI PRISM Big Dye™ Terminator v2.0 Cycle Sequencing Ready Reaction kits were from Perkin Elmer Applied Biosystems Division. QIAamp Blood Kits, QIAamp Blood Maxi Kits, DNeasy Tissue Kit, QIAquick PCR Purification Kits, QIAquick Gel Extraction Kits and QIAprep Spin Miniprep Kits were from QIAGEN Ltd. PCR-Script™ Amp Cloning Kit was from Stratagene. GenomiPhi™ DNA Amplification Kit was from Amersham Biosciences.

2.2.3 Stock Solutions, Buffers and Media

2.2.3.1 General solutions and buffers

TE 10mM Tris, 1mM EDTA (pH 8.0)
TBE buffer 89mM Tris borate, 2.5mM EDTA (pH 8.3)

Where appropriate, solutions and buffers were autoclaved at 105kPa for 20 minutes before use.

2.2.3.2 Gel loading buffers

Orange G	0.25% w/v orange G, 20% w/v Ficoll 400, 100mM EDTA
Blue dextran	Blue dextran/EDTA and deionised formamide in a 1:5 ratio
Genotyping loading mix (ABI 377)	3:1:1 ratio of deionised formamide: GS-500 TAMRA internal lane standard: blue dextran (50mg/ml)/EDTA (50mM)
Genotyping loading mix (MegaBACE)	7.5µl distilled H_20, 0.5µl MegaBACE™ ET size standard

2.2.3.3 Bacterial culture media

LB broth	10g bactotryptone, 5g yeast extract, 10g NaCl (pH 7.0) per litre
LB agar	As above, but with 17g bacto-agar added prior to autoclaving
NZY broth	10g of NZ amine, 5g yeast extract, 5g NaCl (pH 7.5) per litre
NZY^+ broth	9.6ml NZY broth, with 100µl 1M $MgCl_2$, 100µl 1M $MgSO_4$ and 200µl filter-sterilised 20% w/v glucose solution added immediately prior to use

LB broth, LB agar broth and NZY^+ broth were sterilised by autoclaving. Ampicillin (50μg per ml of medium) was used for host cells carrying the pPCR-Script Amp cloning vector and was added to aliquots of LB broth immediately prior to use, and to LB agar once it had cooled after autoclaving, immediately prior to pouring it into sterile petri dishes. LB agar plates were stored at 4°C, wrapped in Clingfilm, for 1 month or less.

2.2.3.4 Lymphblastoid cell culture media

Cell lines were cultured in RPMI 1640 medium, supplemented with 10% v/v FBS, 2mM L-glutamine, 100 IU/ml penicillin and 100 μg/ml streptomycin. For long term storage at -70°C, the medium was further supplemented with 10% v/v DMSO and additional FBS, to give a final concentration of 20% v/v.

2.3 Methods

2.3.1 Preparation of DNA

2.3.1.1 Extraction of DNA from human venous blood

For routine isolation and purification of genomic DNA from whole venous blood, QIAamp Blood kits (QIAGEN Ltd.) were used in accordance with the manufacturer's instructions. These kits were chosen because they are suitable for DNA preparation from fresh and frozen blood, and yield pure DNA, free from contaminants and PCR inhibitors, which can be used directly as templates for PCR amplification. QIAamp DNA Blood Maxi kit was used for extracting DNA from 10ml blood. QIAamp Blood Mini kit permits DNA extraction from as little as 200μl, which is convenient if it is only possible to take a small blood sample, especially from a young child. In this system, white blood cells are lysed in the presence of proteinase K and the DNA is then adsorbed onto a silica membrane by brief centrifugation in a 'spin column'. The salt

and pH conditions ensure that protein and other contaminants are not retained on the membrane. The DNA is washed and then eluted from the column membrane.

2.3.1.2 Extraction of DNA from EBV immortalised lymphoblastoid cell line

Immortalised lymphoblastoid cell lines were established for the two affected girls (VI.3 and VI.5) from Family 2 and the two affected boys (IV.1 and IV.3) from Family 1. Establishing a cell line is convenient because it allows the preparation of as much DNA as needed without having to take repeat blood samples from affected children. Cell lines were cultured in RPMI 1640 medium with glutamax, supplemented with 10% v/v FBS, 100 IU/ml streptomycin at 37°C in a humid atmosphere containing 5% CO_2. Cultures were expanded by 2-3 volumes every 2-3 days until they reached a volume of 50ml. Cells from each culture were harvested, pelleted by centrifugation at 1,000 rpm for 5 minutes and resuspended at $5x10^6$ cells/ml in freeze mix which consists of medium supplemented with 10% v/v DMSO (dimethyl sulfoxide) and additional FBS, to give a final concentration of 20% v/v. 1ml aliquots were then placed in cryotubes, allowed to freeze slowly at -70°C overnight and transferred to liquid nitrogen for long term storage. The remaining cells were pelleted by centrifugation at 1,000 rpm for 5 minutes and resuspended in 10 ml of PBS to wash them, pelleted again at 1,000 rpm for 5 minutes and resuspended in 5 ml of PBS. DNA was extracted with QIAamp Blood Maxi kit, in accordance with the manufacturer's instructions.

2.3.1.3 Extraction of DNA from paraffin embedded tissue

DNA extraction from a post mortem paraffin embedded specimen was required, and for this purpose the QIAamp tissue kit (QIAGEN Ltd.) was used. This kit allows rapid isolation of genomic DNA from fixed tissue using a very similar principle to that described above for DNA extraction from blood. The paraffin was removed from the sample by extraction with xylene and the manufacturer's protocol was followed.

2.3.1.4 GenomiPhi DNA amplification

To amplify small amounts of DNA, the GenomiPhi™ DNA Amplification Kit (Amersham Biosciences) was used in accordance with the manufacturer's instructions. This kit was chosen because it allows DNA amplification from nanogram amounts of starting material. It utilizes bacteriophage Phi29 DNA polymerase to exponentially amplify single- or double-stranded linear DNA templates during an isothermal (30°C) strand displacement reaction. 1µl of template DNA (minimum of 1ng in 1µl) was added to 9µl of sample buffer containing random hexamers that non-specifically prime polymerisation catalysed by Phi29 DNA polymerase. The sample was heated to 95°C for 3 minutes to denature the template DNA. After cooling at 4°C, the sample was mixed with 9µl of reaction buffer containing salts and deoxynucleotides and 1µl of enzyme mix and incubated at 30°C for 16-18 hours. After amplification, the sample was incubated at 65°C for 10 minutes to heat-inactivate the enzyme. Phi29 DNA polymerase starts replication at multiple sites on the linear single-stranded DNA and while polymerisation proceeds, strand displacement of downstream replicated DNA generates new single-stranded DNA. Subsequent priming and strand displacement produce large quantities of double-stranded DNA that can be used directly in other applications without purification.

2.3.1.5 Preparation of plasmid DNA

Preparation of plasmid DNA was carried out using QIAprep Spin Miniprep kits (QIAGEN Ltd.), in accordance with the manufacturer's instructions. In this system, bacterial cells are lysed under alkaline conditions and selective adsorption of plasmid DNA is ensured by the use of silica-gel membrane. The DNA is then washed to remove salts and eluted. This method is fast, simple and yields pure, concentrated DNA that can be used directly in subsequent applications with no need for phenol/chloroform extraction or ethanol precipitation.

2.3.1.6 Quantification of DNA

DNA was quantified by diluting a 10µl aliquot of DNA in 1ml TE buffer and the absorbance of the solution was read at 260 nm using a spectrophotometer (JENWAY 6505 UV/Visual spectrophotometer). The DNA concentration was then determined using the A_{260} value (an A_{260} of 1 corresponds to approximately 50µg/ml for double-stranded DNA).

2.3.2 PCR Amplification of DNA

2.3.2.1 Microsatellite marker sets for genotyping

For the first-pass genome-wide screen in patients IV.1 and IV.3 from Family 1, patients VI.3 and VI.5 from Family 2 and patient II.6 from Family 4, the Single Chromosome Scan linkage marker set (Research Genetics, Inc.), which has an average resolution of 7 cM, was used. This set consists of 516 fluorescent dye-labelled (6-FAM, HEX, and TET) primers for trinucleotide and tetranucleotide microsatellite markers that can be pooled and loaded into a single gel lane. Markers are grouped by chromosome in panels and can be multiplexed by colour and allele size within each panel(Levitt *et al.*, 1998). Gaps larger then 10 cM obtained in the results from the first-pass genome-wide screen were filled in, where possible, using two other linkage marker sets available in the Molecular Medicine Unit: the ABI PRISM™ Linkage Mapping Set (Perkin Elmer Applied Biosystems Division) and the Genome-wide Human Screening Set Version 8a (Research Genetics, Inc.) which both have an average resolution of 10 cM. The ABI PRISM™ set consists of primers to amplify dinucleotide microsatellite repeats while the Research Genetics set consists of primers for trinucleotide and tetranucleotide microsatellite markers.

2.3.2.2 Primer design for genotyping

Where none of the three available sets contained primers to a suitably located marker, new fluorescent primers were designed. The primer sequence was often taken directly

from public databases (uniSTS at http://www.ncbi.nlm.nih.gov/genome/sts (Dib *et al.*, 1996) and UCSC Human Genome Browser database). In some cases, for example when published primers seemed very short or had a low annealing temperature, primers were re-designed to be at least 20-22 bp long and with higher annealing temperatures with the aid of PRIMER3-Design primers pair and probes program at http://workbench.sdsc.edu. The program was used to select primers from the genomic sequence available around the microsatellite repeat. 6-FAM and TET 5' labelling (blue and green respectively) was preferred as HEX-labelled products (yellow) are usually faint and difficult to read.

For haplotype analysis in CLS Families A, B, C and D, fluorescent primers for two markers on chromosome 15q13 were taken from the above linkage sets available in the laboratory, while the others were designed as described above. Details of the markers used are shown in Table 2.1. Details of the primer sequences and optimally determined PCR conditions for the markers used for haplotype analysis in JATD families on chromosomes 4p16, 12p, 6, 16, 3q, 8q24, 12q and 15q are listed in Tables 2.2, 2.3, 2.4, 2.5, 2.6, 2.7 and 2.8. Dye-label, heterozygosities and allele sizes are also given.

2.3.2.3 Primer design for direct sequencing

For sequencing of exons, oligonucleotides primers were designed from genomic sequence with the aid of the program PRIMER3-Design primers pairs and probes as described in section 2.3.2.2. Since these primers were used for direct sequencing of the PCR product after the initial PCR amplification, they were designed to be far enough 5' and 3' from the region of interest so that good quality sequence of the entire exon and splice site could be obtained. Details of the primer sequences and their optimally determined PCR conditions for direct sequencing are shown in Tables 2.9-2.15.

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D15S165	180-208	0.79	Marker from ABI PRISM Linkage Mapping Set	HEX	di	55
D15S976	142-154	0.63	CCATGAGCCACTATGCC GCCTATGACCCAGCAATTC	FAM	di	53
D15S1031	299-321	0.74	CAATCATGTGAGCCAATTCC ACCCTGCATCATCCTCGTT	TET	tetra	53
D15S1010	211-237	0.8	TAGGGGCAAATTCAATCTC TTCACACAGCGTGGAAG	TET	di	53
D15S144	156-170	0.67	AGAGACATCAATACATGATTGGG GAGGTGAAGATTGCCGTGAG	FAM	di	53
D15S1007	165-189	0.87	GGGGAACCTACACTTCCG CCAGGAATCTCAAATGGCTT	TET	di	53
D15S1040	197-211	0.77	TGGGAGGCTGAGTCAC AAAGCCAAATGTAGAGGAAT	FAM	di	53
ACTC	68-92	0.87	Marker from Research Genetics linkage marker set	FAM	tetra	55
D15S118	218-230	0.75	TCAAAGACCCATATCAACCA GTGCTGAAAAGCGACACTTA	TET	di	53

Table 2.1 Microsatellite repeat markers on chromosome 15q used for haplotype analysis in CLS families.

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D4S3360	166-182	/	CTAGCTTTGATTCTATTGACC GGTCTAAATCAATGACCTAAGC	FAM	di	55
D4S412	159-171	0.76	Marker from ABI PRISM Linkage Mapping Set	HEX	di	55
D4S2375	~288	/	GTTGTTTCTCCTTTCATTAGTGC ACCCCCTAACTGTCTCACCT	TET	tetra	55
500H20P5	~235	/	TGCAACCAAAACTAAGAG CATTCCCCAAGATCAACCAGA	FAM	di	59
500H20P3	~165	/	TTTGCTCTACTTTCCTTGTTGTTT GATAGTACCCAGCCTTCA	TET	di	55
D4S2354	~218	/	TCTGCCTTTCTCCCTAACAGAG TTAGGCTAAGTGAGCGTTCCTC	FAM	di	60
***EVC* - Exon 18**	~225	/	AGCCTGCCTGTGGGGACTG GCTGGTGCACTGGGAACTG	FAM	tetra	67
D4S827	~216	/	CAGGGGAAGGAGAGGACTTCAC GGCGTCTTGTAGTTGACAGTTTAC	TET	di	55
D4S2366	120-144	0.77	Marker from Research Genetics linkage marker set	TET	tetra	55
D4S2983	~200	0.87	CCAGTTGGCAGGGATAAAATAG CACAAGTTCAAAGAGGCATCAG	TET	di	60

Table 2.2 Microsatellite repeat markers on chromosome 4p16 used for haplotype analysis across the *EVC/EVC2* locus in JATD families.

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D12S373	208-224	0.76	Marker from Research Genetics linkage marker set	TET	tetra	55
D12S1066	208-224	0.81	Marker from Research Genetics linkage marker set	FAM	tetra	55
D12S1042	118-136	0.81	Marker from Research Genetics linkage marker set	TET	tetra	55
D12S345	208-242	0.87	Marker from ABI PRISM Linkage Mapping Set	TET	di	55

Table 2.3 Microsatellite repeat markers on chromosome 12p11-p12 used for haplotype analysis in JATD families.

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D6S276	197-229	0.83	Marker from ABI PRISM Linkage Mapping Set	FAM	di	55
D6S2427	198-223	0.77	Marker from Research Genetics linkage marker set	FAM	tetra	55
D6S1017	150-174	0.68	Marker from Research Genetics linkage marker set	TET	tetra	55
D6S2410	236-252	0.72	TTGCCTTGTTAACATGTACAGG AGTGCTAGACTGGGAGCAGA	FAM	tetra	55
D6S2436	121-149	0.75	CAGGATTTTTCAGAAAGGCA CTGTCACAGACTTTTTCAGCC	FAM	tetra	60
D6S1581	~238	0.72	TTAATGTCTTCCAGGCTCATCC GAATACCTGCATTCCCACATTT	TET	di	60
D6S1273	188-205	0.8	Marker from Research Genetics linkage marker set	HEX	tetra	55
D6S1719	~180	0.75	ATGGAACTACTCCCATTCCAAC GAGATACAAGCAGGAGGTAGCA	TET	di	58
D6S1027	177-139	0.77	Marker from Research Genetics linkage marker set	TET	tri	55
D16S521	~234	0.71	AGCTGGCAGTGAGCCAAGAT AAATGTCGTATTCACCATGTGC	FAM	di	60
D16S2622	71-87	0.88	Marker from Research Genetics linkage marker set	FAM	tetra	55
D16S423	133-157	0.73	Marker from ABI PRISM Linkage Mapping Set	HEX	di	55

Table 2.4 Microsatellite repeat markers on chromosomes 6p21, 6q25-q27 and 16p13.3 used for haplotype analysis.

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D3S1764	217-261	0.79	Marker from Research Genetics linkage marker set	FAM	tetra	55.5
D3S1309	~171	0.74	CTTTGGGGAATCATTAGTCTGT CTTGCATCCCTGAGATAAATCC	TET	di	60
D3S3694	~299	0.84	TCCATCAACATGGGAATGAATA CCCAACTCCTCACTACTCCAAC	FAM	di	60
D3S1569	147-169	0.8	Marker from ABI PRISM Linkage Mapping Set	FAM	di	55.5
D3S1593	~210	0.78	GGAACAGTATTCTTGGTTGAAAG G	TET	di	60
D3S1744	131-173	0.77	Marker from Research Genetics linkage marker set	TET	tetra	55.5
D3S1306	~166	0.72	TTTTTGTGTTGCACAGTGAAAAT TGGATTTAGTCTGCTGGTAGGAA	FAM	di	57
D3S1555	~163	0.8	TTCCCAGTCACAAACACACTT CATATGCCATTCTTGGATTTCA	FAM	di	60
D3S1308	~196	0.69	ATGCCTAGTCAGTCCAAGGAAC TTTTATTGCTTTTGCCACCTCT	FAM	di	60
D3S3705	~228	0.67	GGCAATTTTGAAATCTCTTCCA TGCGTAGTCAAGATTGTATGGC	TET	di	60
D3S1299	~201	0.62	GGGGGAAAAATTGAAAGACAAT TGCACGCTACCAAACTACTGTT	TET	di	60
D3S1279	264-282	0.85	Marker from ABI PRISM Linkage Mapping Set	FAM	du	55.5
D3S4531	257-273	0.69	Marker from Research Genetics linkage marker set	TET	tetra	55.5
D3S1746	248-284	0.85	CCCCTGCAAACAAATTTTTT GTGTCTTAAAATAATCCAGGCG	TET	tetra	60
D3S1280	~222	0.75	CTGCTGGAGGCTGTTTTACAGT CACCAGAGGACTACAACCCACT	FAM	di	60
D3S1237	~297	0.79	CAGGAAGAGTGATGTGCCTACA GTTAATGCGGCAAAGAGTAACC	FAM	di	60
D3S3710	~243	0.74	TATTCCACCTGACTCTATTGCCC TGCTTTTATGCAAGATGAAGTGG	FAM	di	62
D3S3531	210-256	0.73	AAGAGCCTAAATTAGGAAAATAA AA AAACAAAGCAAGACTCCATC	TET	di	60
D3S1607	~235	0.8	TTTGTAATAAACAGCAGCCTGTC TGCACGTAACGAAAATGAAGAG	TET	di	60
D3S3692	~242	0.62	TTTTGCAAGAGCCTGTCACTAA CATCTTGGGAATGGACTACACA	FAM	di	60
D3S3575	~192	0.79	GGGAACAGACAATTCCACAGAT TCAGGGGTTGCTTAAGAGACTT	FAM	di	60
D3S3580	~242	0.66	GGCCATTGCAGAGAAGTTTGAT TTGGGGCTACAGTTACAAGGGT	FAM	di	62
AC026118	~183		TATAAGCGAGGGTGACTGCTT GAAGCTCATTGCCTTTCAAAC	FAM	di	59
AC024221	~237		AACAAAATGGAATCACAGTTGC GCAAATAGCAAAATGGTTCTGA	TET	di	59

AC069224	~223		ACTCTGGCCCTTAAATGGATAA AGGTCTATTTTTCAACAGGGGA	FAM	di	59
D3S3579	~183	0.79	TATGGGAAATTTTGAGCAACCT TGGGGATTTGACAGTATTAGGG	TET	di	60
D3S1268	~209	0.86	GCTTAGTGAAAAGGTGCAGTGGT CTTGGCCAATTCCACAACTTTAG	TET	di	62
D3S3668	~225	0.83	TGGGAATTAAAAACTTCAGGAGA TGAGCTCACTGGCTAATCTCAA	TET	di	60
D3S1763	260-280	0.8	Marker from Research Genetics linkage marker set	FAM	tetra	55.5
D3S1614	100-122	0.83	Marker from ABI PRISM Linkage Mapping Set	TET	di	55.5
D3S1282	~167	0.78	GCTGTGATTTGTTTCTTCTGAGAT TTCCACTGTATCACAGGTTTTCA	FAM	di	57
D3S3523	~193	0.68	TGAAAACAAAATCAGCAGACTTA CA CCAACCTTGCCACTATCATGT	TET	di	60
D3S3723	~166	0.63	TAAGTGAATTTGTCAGGCACCA TGGACCATGACAGAGTGGTAAG	FAM	di	60
R1 AC026316	~276	/	AAGTAGGGTGGCCAATGTTGTT TTCCTCTCAGCATGGTCTTCCT	FAM	di	62
R2 AC061708	~199	/	CCTCCCAACTCAGGATTTCAAC TGCTGGAAGAAGCATATCAGGA	TET	di	62
R3 AC092919	~286	/	CAGGTGGTTGAGAGACCATCAG GTGGATGCCAAATGAAGAATTG	FAM	di	62
D3S1574	~129	0.79	GTTAACTCCCCTACCCCAGTGT TCTTCACAATTTCCATTCCCTTA	FAM	di	60
D3S3053	226-238	0.73	Marker from Research Genetics linkage marker set	TET	tetra	55.5
D3S3725	~204	0.84	CTTAAGGTTAGAGGTGCCATGC TCACAGTGCCCTGACTTTACTG	FAM	di	60
D3S1565	175-189	0.63	Marker from ABI PRISM Linkage Mapping Set	FAM	di	55.5
D3S2427	216-260	0.85	Marker from Research Genetics linkage marker set	HEX	tetra	55.5
D3S3730	~287	0.83	TTGTAATGGCAAGTTGGTTCTG TGAAGATGAGTCCTGAGCATGT	TET	di	60
D3S1262	107-129	0.8	Marker from ABI PRISM Linkage Mapping Set	FAM	di	55.5
D3S2436	164-180	0.67	Marker from Research Genetics linkage marker set	TET	tetra	55.5
D3S1580	212-232	0.84	Marker from ABI PRISM Linkage Mapping Set	FAM	di	55.5
D3S2398	266-298	0.77	Marker from Research Genetics linkage marker set	TET	tetra	53

Table 2.5 Microsatellite repeat markers on chromosome 3q24-q26 used for haplotype analysis

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D8S1128	240-268	0.76	Marker from Research Genetics linkage marker set	HEX	tetra	55.5
D8S1720	~153	0.82	CTGAAAGAAAATGCCCACCCTA GTTTTTGTCTGCCTTTGCTTCT	FAM	di	60
D8S1732	~206	0.78	ATGGGGATACCCCCATGTTTC ATCTGCAAGGGCTCTATTTCCA	FAM	di	62
D8S1701	~275	0.7	TACTTGGTCTTGGCTTTCCAAT AGTGAAACCCCATCTCTACCAA	FAM	di	60
AC131568	~188	/	GAATTTCTGACCAGTTGGCTTC GAGACTGCAGTGAGCTGTGATT	TET	di	60
AC103725	~226	/	ACTGTTAAGTGCTGGAATGGGT ATGAAGACCTGTGTCCCTCCTA	FAM	di	60
AC009682	~232	/	CCATCCCACTACTCTAACCCAAC GGTGACAGAGTAAGATCCCGTCT	TET	di	61
D8S284	268-300	0.83	Marker from ABI PRISM Linkage Mapping Set	HEX	di	55.5
D8S1712	~231	0.72	ACTGACTGGAGACGTGAAATGGT GGATCTCCTCTCTCCTGGAACTC	FAM	di	62
D8S557	~214	0.76	CAGACATGAAATTGCTGAGACA CCTGGGGTCCTAGAGATTTAGA	TET	di	58
D8S558	~164	0.86	TTTTTAAGCCCCTAGTGTGGAA CTGAGAAGTCTCTGGGGCTTTA	TET	di	60
D8S529	~247	0.81	GTAATTAGCATGCTGGCATTGA CATAATCCCTCCCAATCAAAGA	TET	di	60
D8S256	~221	0.83	GGCACTTTGCTTTTCTTCCTGT ATGGAGTGAAGGCTCTCCTCTG	FAM	di	61
D8S1708	~154	0.68	TCACCAAACTTCTAAATGAAGGC GTGAACTGCTGTGTTTATGTGGT	FAM	di	59
D8S1796	~221	0.66	GAGAGCACCTAAGCCAGTGTTT GAGAGCACCTAAGCCAGTGTTT	FAM	di	59
D8S1746	~203	0.7	GGGTTCCTCATCATCAGCTTAG AGATCCAGTGAACATCAACCCT	TET	di	59
D8S1462	151-170	0.75	Marker from Research Genetics linkage marker set	FAM	tetra	55.5
D8S1710	~260	0.78	TATACAGACCTCCATTCATGCG TCAAATGGGCTTTGTTTTTCTT	TET	di	58
D8S537	~182	0.89	TCTATGAATTTTGGACTTGCCA TTTGGTCTCGTATATGCCTCTG	FAM	di	58
D8S1108	~258	/	AATGCCTACAAAAACCAGTGCT GCTTGCCTGGTCTCTTACACTT	TET	tetra	60
D8S534	~177	0.81	CTTGCAGCACGACTCACAATAAC CCATGTTGTCAAAAATGGCAGTA	FAM	di	61
D8S1100	183-192	0.65	CTGTGATTATTGCAAACCCC CCAGAGTACCTGGCATATGG	TET	tri	55.5
D8S1783	~172	0.78	TTATCAGAAAAGATGATGAAGGCA ATAATATCCAGAAAATCAAGGCCA	FAM	di	60
D8S1761	~287	0.71	TATCCAATGTTGATGGCACCTC TGTTATCTCTGGCCTGTTGGA	FAM	di	60

D8S274	~117	0.77	ATAATCCACTAAGCCCAGCAGA TTCTAAGTTCCCCAATTTGGAC	TET	di	60
D8S272	~239	0.8	CAAAATGCTATTGAGTTCTTCCCT GCAACCTGGATCTTATTTGCTTAT	TET	di	61
D8S1837	~230	0.81	CTAATGAAAGGCTGACCTCCTG TTATGCTCTGCTCTCTGCTTTG	TET	di	60
AFM316xb1	~192	/	GATTTCCAAGAGAAATTGGCAC TCCTTCTTTGGAAGGATGTGAT	TET	di	60
AC087711	~198	/	ACTGGAATGATACCACCTGCACT GTTCTCCAGAGAAACAGGACCAA	TET	di	62
AC087711-2	~244	/	TGAGCTGGGCTATGACTTTACA AATGACCCTTGAAAGGAGATGA	TET	di	60
AC068467	~212	/	TATCCCAAGCATGGAGCTAGTGT TTAGTAAGGCCTTGGAGGAAAGC	FAM	di	62
AC105130-R1	~257	/	ACTCACCCCTGGTGATTCTGAT ATTTCATATTTGGGTGATGGGC	FAM	di	60
AC053480-R1	~286	/	CAGAGAATAGGATGGTGGCTTC CATCACCTCCAAAATTTCCTGT	FAM	di	61
D8S1743	~245	0.83	TGGTTAATAAAGGGCAGGGAAA AGTTTCCTCCCACTCCCCTTAC	TET	di	60
D8S373	~271	0.78	CTTGGTCTCCTCTGATGGGTAG CAACAAGTAAGTGCTGGCTGTC	TET	di	60
D8S2334	~155	/	CTGAAGCCTAATGCATCCAATG TTCCAGCATAGCACTTCTGGTC	FAM	di	61
D8S1926	~253	/	TGAACTTGAAAAATATTTGGGCTT TGATCATGTGGGACAGATAAACTT	FAM	di	60

Table 2.6 Microsatellite repeat markers on chromosome 8q24 used for haplotype analysis.

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D12S346	184-208	0.84	Marker from ABI PRISM Linkage Mapping Set	TET	di	55
PAH	229-257	0.8	Marker from Research Genetics linkage marker set	HEX	tetra	53
D12S78	171-201	0.91	/	FAM	di	55
D12S317	250-280	0.71	/	HEX	di	55
D12S1342	266-288	0.83	/	FAM	di	55
D12S353	89-105	0.77	/	HEX	di	55
D12S1605	195-205	0.78	/	FAM	di	55
D12S84	199-219	0.84	/	HEX	di	55
D12S1583	219-247	0.87	/	FAM	di	55
D12S1645	212-248	0.77	/	HEX	di	55
D12S1339	266-278	0.71	/	FAM	di	55
D12S1646	247-259	0.72	/	FAM	di	55
D12S1341	203-221	0.79	/	FAM	di	55
D12S79	155-181	0.87	Marker from ABI PRISM Linkage Mapping Set	TET	di	55
D12S2070	86-104	0.79	Marker from Research Genetics linkage marker set	TET	tetra	53
D12S395	181-217	0.85	Marker from Research Genetics linkage marker set	HEX	tetra	53
D12S324	230-252	0.69	Marker from ABI PRISM Linkage Mapping Set	TET	di	55

Table 2.7 Microsatellite repeat markers on chromosome 12q23-q24 used for haplotype analysis. / means primer sequence is unknown.

Markers	Allele size range (bp)	Het. Index	Primer sequences forward and reverse (5' to 3')	Labelling dye	Repeats	Annealing temperature (°C)
D15S817	140-164	0.79	Marker from Research Genetics linkage marker set	HEX	tetra	55
D15S1021	~221	0.79	CTACAGCCTGGATGACAGAGT GGCATCTCTTGCCATTTTCTTA	TET	di	60
D15S128	193-211	0.78	Marker from ABI PRISM Linkage Mapping Set	TET	di	55
D15S822	242-306	0.86	/	FAM	tetra	55
D15S1002	~136	0.79	GTATCCCAAGGCCATACCCTG GTCTTGCTAGGTTCCCCTTTTT	TET	di	62
D15S219	203-227	/	GAGTCTGTCTGTGGGAGGGAG AATGCTTTCTAAGCCAGCCAA	FAM	di	62
D15S156	~251	0.51	GACACAAAAGCCAGGTATGACA TGTGGATTTATTTCTGGGCTCT	TET	di	60
D15S217	211-250	/	Marker from Research Genetics linkage marker set	FAM	tetra	55
AC079090	~237	/	GTGTGCCTCTCTGTACACCTGG TTCATCGGTCACCAAGTCCAT	TET	di	62
AC024474	205-215	/	CACATAAAACTTCAAAGGCCCTG GATATGCTGCCTTGAAAGTGTCC	FAM	di	62
D15S1019	~222	0.63	ATTCTGGACCACGCATACTAGG TTTTTCCCACTCAATATCAGGC	FAM	di	60
AC090763	~267	/	ATTTGGCTCCTGCTTCATCTTC GCTCAACTAGCACCCTGGTTTT	TET	di	62
D15S1048	203-233	0.84	/	TET	di	55
AC022613	~274	/	AGGTTGCAGTGAGCCAAGACTG TTCCTTCAGTCCCATAGCATCA	FAM	di	62
D15S1043	~231	0.52	AAGAACTATGAGGCAAGGAACG TAGGGGGTTCAGTGTGCTTATT	TET	di	59
D15S165	180-208	0.79	Marker from ABI PRISM Linkage Mapping Set	HEX	di	55
D15S976	142-154	0.63	CCATGAGCCACTATGCC GCCTATGACCCAGCAATTC	FAM	di	53
D15S1013	169-179	0.53	/	TET	di	55
D15S1031	299-321	0.74	CAATCATGTGAGCCAATTCC ACCCTGCATCATCCTCGTT	TET	tetra	53
D15S1010	211-237	0.8	TAGGGGCAAATTCAATCTC TTCACACAGCGTGGAAG	TET	di	53
D15S231	102-114	0.5	Marker from Research Genetics linkage marker set	FAM	tri	55
AC019278	~196	/	ATCCTGATGATGCCACTAAGCA ATGTGTTTCCCGCTCTCAATTT	TET	di	62
D15S995	~242	0.67	CAAAGCAAAGAAGGCATAGTTG CTGGCTCTAAGCTCATCTCTCA	FAM	di	59
D15S1007	165-189	0.87	GGGGAACCTACACTTCCG CCAGGAATCTCAAATGGCTT	TET	di	53
AC087638	~295	/	TGTTCGCACTCTAGCTTTCCT GAAGAGGTGGTTGGTGATGATG	FAM	di	62
D15S1232	~309	/	TGTCCATGCTCTTAACCATTCTT ATGATGTTGCTCCACTCTTTTCT	TET	di	60

ACTC	68-92	0.87	Marker from Research Genetics linkage marker set	FAM	tetra	53
D15S971	~298	0.82	CCACCACAGCTTGCTACTTAACC AACAGAGCGAGACTCTATCTCGAA	TET	di	62
D15S1042	~278	0.82	GGCTAATGCCATGATAAGCAGA GCCAAGTTTTCCTCATCCATCT	FAM	di	62
GATA50C03	274-290	0.72	Marker from Research Genetics linkage marker set	FAM	tetra	53
D15S994	~183	0.72	CATCCAGGAAGTAGTGCAGGTG AGTGCATCTCCTAAGGGCAAAG	FAM	di	62
D15S659	174-210	0.84	/	TET	tetra	55

Table 2.8 Microsatellite repeat markers on chromosome 15q13 used for haplotype analysis. / means primer sequence is unknown.

Primer pair	Primer sequences forward and reverse (5' to 3')	Annealing temperature (°C)	Product size (bp)
FMN 1F *FMN 1R*	ATCAGTTTACTGCTCACCTCGG CTCATTCCAAGAACCCTGAAAG	60	294
FMN 2AF *FMN 2AR*	ATCTGGCTGTTGTCTGATTCCT GAAAAATATGTCGCCTGGATGT	60	498
FMN 2BF *FMN 2BR*	AGGAGTCAGACATCATCAGCCT CCAAAGGTTAGGCCTTGTTCTA	60	473
FMN 2CF *FMN 2CR*	GAAGAGGACCAAAAGAAAAGGG CTTTGGAGACAGGTTTCTGCTT	60	494
FMN 2DF *FMN 2DR*	AGAGGCAGAAAAGGATGAGATG TCCTCTTGTTTCTCGTTTCTGG	60	458
FMN 2EF *FMN 2ER*	CACACGAGTGTCCCTCACAC GCAGACTGATAGTTCCCCAATC	60	600
FMN ex2FF *FMN ex2FR*	AGCCTCCATTTCTAGTGTGTCG TTAAGAGCAGCTTCCCTCTCAC	60	447
FMN 3F *FMN 3R*	GCTATAAAGTGGCGAAACCTGT CAATGAACACCCTTCTCTCTCC	60	365
FMN 4F *FMN 4R*	TTTGCATTCTGTCTGCTTCTGT CAGCTGACCACGGTTAATACAA	60	194
FMN 5F *FMN 5R*	GAGTTCCCCAAATATGGCTTTA CACAGCACAATCACCAAGATTT	60	283
FMN 6AF *FMN 6AR*	GAAGTTTCAGAAGAGCAGCCTC CTATGGAATGCTTTCAGCACAG	60	483
FMN 6BF *FMN 6BR*	AATCCAGAAGAGGAAGATGCTG ATCCTTTGAGGATCTTCTGCAC	60	482
FMN 6CF *FMN 6CR*	GGAAAATACTGTCACTGGGGAA GTCAATGTTGAGCAGCAGAGAG	60	493
FMN 6DF *FMN 6DR*	GAAGAAAGCAACTGATCCGAAG TAGGATTCACTTTGGGAGGAAA	60	472
FMN 7F *FMN 7R*	CATTTTTCAGGAATTAATCGGC AACCCCAAATTCTGAGAGTCAA	60	273
FMN 8F *FMN 8R*	TGATGTCAAATTGTGGCATTTAT TTCAGCATAGGAGAAAACCAAG	59	242
FMN 9AF *FMN 9AR*	ATGCTACAGTGGCACCATGA GGAGACAAAGATCCAAGTCCTG	60	545
FMN 9BF *FMN 9BR*	CAAATGACCACAAAGACATCCA AATATCACTTGGGCCAAATCC	60	519
FMN 10F *FMN 10SR*	CATTCCTTCCTGTTTTGGTTCT CAGCTTTGCCATAATCACTCAG	59	228
FMN 10F *FMN 10LR*	CATTCCTTCCTGTTTTGGTTCT TCATGAACTATTGGCTCCCTTT	59	489
FMN 11F *FMN 11R*	AAAGACTGCATGGTCATTTGAG TGAGTTGATGGTTCTGTAGCTGA	60	285
FMN 12F *FMN 12R*	GGAAGCTGGATGTGGAGTAAAC GCCACGTCAAAACCCTTACTAC	60	246
FMN 13F *FMN 13R*	TGCTGCTAGTCTAGGAACCACA TTTAGCTGGCTGCACATTTAGA	60	308
FMN 14F *FMN 14R*	CCCTGGGAATCCCTTGAGTCTA GCATTAAAGTGGTCTCTCCCAC	60	361
FMN 15F *FMN 15R*	GTTACTGAATGAACATGGGTGC AATAATGGCAGATCATGGGAAC	60	254

FMN 16F *FMN 16R*	AGATTTCACCTCTCAACCGTGT ATTCTATGTTGGGCTTGCAGTT	60	243
FMN 18F *FMN 18R*	TCAACTGAAAGTGACAATTGGG TCCCACTCTATGAACCAGACCT	60	272
FMN 19F *FMN 19R*	CACAATTGATAGCCAAGTGCA ATGGAATGAAGGATGTTTCAGG	60	246
FMN 21F *FMN 21R*	TTACGCAAAGGAGACAACAGAA TGGGGAAGTAAAAATGTTTTGG	60	227
FMN 22F *FMN 22R*	TGGCAATTTCTCACCTCTACCT CTTCCCCTATGACCACTTCATC	60	329
FMN 23F *FMN 23R*	TTGGGTTTATGTTTAACCTGCC CATGACAATCGTTAGATTATGTTGAA	60	224
FMN 24F *FMN 24R*	ATTCCGGAAACTCAGCATTCTA GGCTTCCACAAGAAACAGTCAT	60	337

Table 2.9 Oligonucleotide primer pairs used for PCR amplification and direct sequencing of *FMN-1* exons.

Primer pair	Primer sequences forward and reverse (5' to 3')	Annealing temperature (°C)	Product size (bp)
GRE 1F *GRE 1R*	ATTTAAACGGGAGACGGCG GCGATGGTTCTTCACAATTCAC	62	341
GRE 2AF *GRE 2AR*	TGCGTTTAAATGCTAGGTGCTA GTAACAGAAGCGGTTGATGATG	60	448
GRE 2BF *GRE 2BR*	GGGAGGAGGTGCTGGAGTC GGGTTCTTCTGGCCTCTAGGTT	62	455

Table 2.10 Oligonucleotide primer pairs used for PCR amplification and direct sequencing of *GREMLIN* exons.

Primer pair	Primer sequences forward and reverse (5' to 3')	Annealing temperature (°C)	Product size (bp)
SHOX2 1AF *SHOX2 1BR*	TCTGCTGGCAGAGGTTGAGC CCCCACCCTTCCACTATCACT	62	571
SHOX2 2F *SHOX2 2R*	ACAGGATTTGCTGTGCTGTTTT CCAGACACTAGAAGCACCACCT	62	408
SHOX2 3F *SHOX2 3R*	TGCTTGCTGTATCTCCCAATTC GGACTCCATTAACAGCCTTTGA	62	326
SHOX2 4F *SHOX2 4R*	ACCCTTTGGAACCCTGAAAAAT CAGAAGGGACTAGGCATTCAGG	62	356
SHOX2 5F *SHOX25R*	CCCAAACACAACCCAACTCTCT TCTCAAAGGGGTAACGGAGAAG	62	462

Table 2.11 Oligonucleotide primer pairs used for PCR amplification and direct sequencing of *SHOX2* exons.

Primer pair	Primer sequences forward and reverse (5' to 3')	Annealing temperature (°C)	Product size (bp)
WISP1 1F *WISP1 1R*	GAAAATGCAGGGTTTGTCCTTC GCTCTGTCCTTGCCCTGTAAGT	62	394
WISP1 2F *WISP1 2R*	GAACAGCATGAGGACAGGAATG TGTATCTCCTGCTGAACGGAAG	62	499
WISP1 3F *WISP1 3R*	CTTTGTGCCTCTGTTCCTCCTC AGGTAGTGGACTGGGGAATGAG	62	506
WISP1 4F *WISP1 4R*	CCTGGGTACCTTTGCTGGTTAG TTCCCACTTTACAAATGCTCCC	62	466
WISP1 5F *WISP1 5R*	AGGTGGAATGCTCCCACATAGT AGAGACAGAAATGGAGGCCAAG	62	540

Table 2.12 Oligonucleotide primer pairs used for PCR amplification and direct sequencing of *WISP1* exons.

Primer pair	Primer sequences forward and reverse (5' to 3')	Annealing temperature (°C)	Product size (bp)
GLYLU BF *GLYLU BR*	GATTGGTTTCCTGTTGTAAAAGCC ACTTGGCAAGAAAACACAATGAAA	60	327
GLYLU B2F *GLYLU B2R*	ATGAAGCTTGCCTGTGTAGCTG TTTGCGCTGTTAATTGGACACT	62	418
GLYLU CF *GLYLU CR*	CTGATTTCCTATGGCTTTCAGGAT CAACATTCACCACAGATTGACTGA	60	302
GLYLU 1F *GLYLU 1R*	TCCGCCCTATATTCTGTGACCT CTCATGATGACCCAGAGTCCAC	62	399
GLYLU 2F *GLYLU 2R*	TTTTAAGCTTTGGAAATTGGAAGG AGCCATGATAGGGAACAGCTACTT	60	320
GLYLU 3F *GLYLU3R*	CATGCATGGAAATTCTGCTTTC TTCAAGGCTATGAGGCTATCCC	60	304
GLYLU 4F *GLYLU 4R*	TGGAGAGGCTTAATACAGACATGC AAACCTGCAGCCTTAGAGAAACTC	60	383
GLYLU 5F *GLYLU 5R*	ATGACTTATCCCAAATGCTTCACA GAGGAAGGGTCAAATTAGCAAAAA	60	361
GLYLU 6F *GLYLU 6R*	TGAACCTCCAAACATTTTGTGAAT ACCAGGGTTATTCCTGTGACATTT	60	379
GLYLU 7F *GLYLU 7R*	GTCCTCAGCAATGTGTTTGTCTCT TTTTTACAGCAAGCAAGCAACTG	61	260
GLYLU 8F *GLYLU 8R*	ACTCTAAACCTTCACTGGGGGTC CACACTCTGTATGGGAGGATGTG	61	363
GLYLU 9F *GLYLU 9R*	AAAGACTTCAGGAAGAAGACTGCTC GTAGAGACACGAAAGCACGAAGAT	60	355
GLYLU 10F *GLYLU 10R*	ATTCAGAACCTTCAGGTCAGGG ATGCAGACAGAAAGTGCAGAGG	62	348
GLYLU 11F *GLYLU 11R*	AAAAGGATGATGGGCAGACAAT TTTGGTCTCACTTTTGCTGGAA	62	387
GLYLU 12AF *GLYLU 12AR*	ATCTTCGAACCGAGTCCACATAG CCTGCTTTGTTTTGAAATTCACC	60	392
GLYLU 12BF *GLYLU12BR*	CATCCCAAATGGATTCTGATGA GGTGGCTTCCACCTACTACCAC	61	405
GLYLU 12CF *GLYLU 12CR*	TATCCAGTGGCAGATGTGGATACT CTCCTGCTCTTCTGTGTGAGAGTC	62	418
GLYLU 12DF *GLYLU 12DR*	TAAGCTTGAAACTCACCCCCTC TTGGCTCTTGCTGTGCATATCT	61	527
GLYLU 13AF *GLYLU 13AR*	CTGTATCTGCTCCCCACATCAG TCCTGGGGTTTTCAAGTGCTAT	61	369
GLYLU 13BF *GLYLU 13BR*	AAAGGGAAGCAGTTTGATGCTC CCTGCTTTATGTTTAGCCTCCG	61	398
GLYLU 13CF *GLYLU 13CR*	GCATTATGGATGACACAGCATTTA TTTGAAACAGTATGCATGAAATGG	61	486
GLYLU 14F *GLYLU 14R*	GAAGTACACAGCTTGGCTTGTGAT TGTCATTCCACAGTTGCTAATTGA	61	332
GLYLU 15F *GLYLU 15R*	AAAACCCAGTAGAAATGGGGAAA CCAGCATCTCTTTTCAAGGTCTC	61	325

GLYLU 16F *GLYLU 16R*	GTAGGCGCCACAGTGTTTAGTG CATGGCAAGAAACAAGATTCCA	62	508
GLYLU 17F *GLYLU 17R*	CTTTGCCAGTTCCCCTACATTC GGGTTCCAAGTGGAAGTACCTG	62	367
GLYLU 18F *GLYLU 18R*	CGTTCATCTCACCAAAATTCCA TGACTAGGCCCTTCTGTGAGTG	61	305
GLYLU 19F *GLYLU 19R*	CAAATTTCCAGAGGGTGTTGAAG TGTTCCCAATAGGCAGTCAATTC	61	268
GLYLU 20F *GLYLU 20R*	CTGCTGTATAACAGGTCTGCCC GCTTCATTCTGAAATGGTGAGG	61	417

Table 2.13 Oligonucleotide primer pairs used for PCR amplification and direct sequencing of *GLYLU* exons.

Primer pair	Primer sequences forward and reverse (5' to 3')	Annealing temperature (°C)	Product size (bp)
ZNF406 1F *ZNF406 1R*	ATCCGTGAAGTGGGATAAAAAGCA GCACTGCTTCCCGACTCGAC	61	438
ZNF406 2F *ZNF406 2R*	GCTGTGGCTGTATTGGAGAGAA AGTTGCTATAGGCGACCAGGAG	61	306
ZNF406 3F *ZNF406 3R*	CACAATCTGGTAGAGGCTGGTG TTTCTAGGCTTCTGCTATGCCC	61	387
ZNF406 4F *ZNF406 4R*	ACACTTGGAGGGTGATTGTGAA ACTTTTTCCAGATGGGTGAGAGC	61	459
ZNF406 5F *ZNF406 5R*	CTACAGTGATTCCGTGGTCTGC GCACCAGCTCTGTGTTCTGACT	61	429
ZNF406 6F *ZNF406 6R*	TCTTGCGTGGGTATCAGACAGT AACAAGCATGCTGATCCTTCCT	61	358
ZNF406 7AF *ZNF406 7AR*	GCGTCTGTCCAGTTTTGTTCAC ACCTTCTTGTCCTGTGGGTCAT	61	484
ZNF406 7BF *ZNF406 7BR*	CGTCAAGAACCTCATCAAGCAC AGCTCTGGTTGATGGAACTGGT	61	398
ZNF406 7CF *ZNF406 7CR*	AAGAAGTTCGTCAGCTCCATCAG GAAACCACCTCTTGAGAATGCAG	61	389
ZNF406 7DF *ZNF4067DR*	GCTACAGCTGGTGGAAGAGGAG AGCTGGGTTTGACCTGAGACAC	61	493
ZNF406 7EF *ZNF406 7ER*	ATCACACTACTGCTGTCCAAGGC GGAGGATGGCTCACTTAGAAGGT	61	471
ZNF406 8F *ZNF406 8R*	CAAGCGCAAGAGGGAGATAGAT CCTGCAGCAGAGACACCTACAT	60	319
ZNF406 9F *ZNF406 9R*	GATGACTGTTGATAGCGTGAGGT GATGCATCTGAAGACCAGATAA	55	426
ZNF406 10F *ZNF406 10R*	GGTGCTGCTTAGCAATGAGGTT CTCCCAAAATGAAGGATGGAAG	60	300
ZNF406 11F *ZNF406 11R*	CTGCTGTGTTCAAAACGGATGT TTCCAGAAGCTCCTTCTAACGG	60	442
ZNF406 12F *ZNF406 12R*	ACCTTTGTGTACCTGGCTCCAT CCAAAGCGAAGGTAAAACCAAC	61	294
ZNF406 13F *ZNF406 13R*	CTTTCAGGGTGTCAGATGCTCA AGCAGAACTGAACTGCAGGATG	61	331
ZNF406 14F *ZNF406 14R*	TTGAAGCATTTGAACAGGAGGA AACGTTTCAGAAGTTCCCCAAA	60	339
ZNF406 15F *ZNF406 15R*	TGCCAGGTAAATGAGATGTTGC ATCTGCAAACGCATTCATTCAG	60	334
ZNF406 16F *ZNF406 16R*	AAACTGAGCGATTACGGAGAGC CAGGGAGAAGGAGAGAACCTGA	60	335
ZNF406 17AF *ZNF406 17AR*	GCACTGGTGCCTCCTAAACTGT GGAGAGTCCTATCAGGCTGCTG	60	477
ZNF406 17BF *ZNF406 17BR*	TCTCCTAGTCCAACTTGGGGTG CTTAATACTCGGGGAAGTGGGAC	60	402

Table 2.14 Oligonucleotide primer pairs used for PCR amplification and direct sequencing of *ZNF406* exons.

Primer pair	Primer sequences forward and reverse (5' to 3')	Annealing temperature (°C)	Product size (bp)
TTC7B 1AF *TTC7B 1AR*	CTGTCTCCCGCCCTAGCCCC GTACCGTTGGCGATGAGCTTGG	68	413
TTC7B 1BF *TTC7B 1BR*	GATGGCGACCAAGAAGGCAGG GACACCCCAGACTCCGCAGAAG	68	393
TTC7B 2F *TTC7B 2R*	GGTGCCCACTTAGACCCTATGA TAAGAAGACAAGCCCTGGCAGT	62	437
TTC7B 3F *TTC7B 3R*	TTTGGGTTGTTTCCACTGTTTG GTCAAATAGGCAGCTCCGAAGA	62	417
TTC7B 4F *TTC7B 4R*	AAGCATCCAGATTGAAAAGGAGG GCTTTGCACACAGAGACAGACAA	62	395
TTC7B 5F *TTC7B 5R*	CCTCTGCACTGATATTTGCCTG GGAAAGCAAGAGAGCCAGGTAG	62	438
TTC7B 6F *TTC7B 6R*	ATGGCTAGGCATGTGCAGAGTA CCTATGAGACAAGTGCCCTGCT	62	274
TTC7B 7AF *TTC7B 7AR*	TATGAGCCTTGCAGTAGTGCTT TGACACTTGTTCATTTGATTTTCTT	59	417
TTC7B 8F *TTC7B 8R*	CCCTTCTGCCTACCACTAAAA TCCCATCCAGGTATACCCTCT	59	405
TTC7B 9F *TTC7B 9R*	ACAGGCCTCTGCTCATTCCTAC AAGATCACATGGCTGCCACTAA	62	393
TTC7B 10F *TTC7B 10R*	CAAGGGAAGCAGATTTTATGGG CACTGAGAACGGACCACGATAG	62	344
TTC7B 11F *TTC7B 11R*	AGCATTCCTGCTGGCTTTTAAC AATCACCCTCGCTTTCTCTCTG	62	418
TTC7B 12F *TTC7B 12R*	GTGGGATTGGGAGATTTGAGAG TCAGACCAAGCAAAGAGGACAG	62	422
TTC7B 13F *TTC7B 13R*	GCACGTAGGCACTCAGACATTG TCAGGACGAAAGACTCTCAGGG	63	309
TTC7B 14F *TTC7B 14R*	TGCAAATGTGGAGAAGACTTGC GAGCCTGGAACTCCTGTGTCTT	62	365
TTC7B 15F *TTC7B 15R*	TGTGAATGAGGCACAATGTGAA ATTATTGACTGCCAGGGTGTCC	62	423
TTC7B 16F *TTC7B 16R*	TGAGTACAGCTGCTTCTGGGTC TCCACTGCTATCATCTGCCACT	62	366
TTC7B 17F *TTC7B 17R*	CATTGGCCTTCAGGAGTCAGTA TCCCCACTGTGAAGAAACACTT	61	372
TTC7B 18F *TTC7B 18R*	GTGCCACTCTTGAATGGAAAT TTCTCAGCACTCAACAACCAC	59	297
TTC7B 19F *TTC7B 19R*	TTAGAAACATCATGGGGGAAGG CCGTGAAGCTTTGATCTCTCCT	62	398
TTC7B 20F *TTC7B 20R*	AGGGAACACGAACATCAGAGTG AGGGAACACGAACATCAGAGTG	62	390
TTC7B 21F *TTC7B 21R*	GAACTGACTGGAAACTCGAGGA GTTGGTTTGGTTGGTTCACTGT	60	459

Table 2.15 Oligonucleotide primer pairs used for PCR amplification and direct sequencing of *TTC7B* exons.

2.3.2.4 PCR amplification of DNA for microsatellite analysis

PCR amplification of microsatellite repeats was carried out in accordance with standard protocols. AmpliTaq Gold (Perkin Elmer Applied Biosystems) was used. This hot start enzyme has the advantage of improving amplification by lowering non-specific background and non-specific primer annealing, as the enzyme is inactive at room temperature. Amplification reactions contained 1x GeneAmp PCR Buffer II, 2.5mM $MgCl_2$, 0.25mM of each dNTP, 1μM of each primer, 0.3U AmpliTaq Gold and approximately 100ng template DNA in a final volume of 10μl. (10x GeneAmp PCR buffer II contains 100mM Tris-HCl (pH 8.3) and 500mM KCl). Initial incubation (to activate the enzyme) and denaturation was at 95°C for 12 minutes, followed by 10 cycles of denaturation at 94°C for 15 seconds, annealing at 53-55.5°C for 15 seconds and extension at 72°C for 30 seconds, and a further 20 cycles of denaturation at 89°C for 15 seconds, annealing at 53-55.5°C for 15 seconds and extension at 72°C for 30 seconds. For the genome-wide screen, individual PCR reactions were done in polycarbonate 96 well microplates or strip tubes and placed in an Eppendorf Mastercycler gradient thermocycler (Eppendorf). For each patient, the PCR products generated by each panel of markers were then pooled. In this way, the products were diluted approximately 1:12 (6-FAM and TET) and 1:6 (HEX).

When filling gaps with single markers, PCR reactions were performed in single 0.2ml Micro-tubes with a domed cap (ABgene) using a T3 Biometra Thermocycler. To assess the success with which the repeats had been amplified, a 5μl aliquot of each completed PCR reaction was first electrophoresed on a 2% NuSieve GTG agarose gel. Depending on the quantity of the product obtained, the remaining sample was diluted 1 in 5 or 1 in 10 in sterile purified water.

2.3.2.5 PCR amplification of DNA for sequencing

PCR amplification of *FMN-1* and *GREMLIN* was carried out in accordance with standard protocols. Biopro™ *Taq* DNA polymerase (Bioline) was used except for

FMN-1 exon 1 and *GREMLIN* exon 1. Amplification reactions contained 1x NH_4 PCR buffer, 1.5mM $MgCl_2$, 0.25mM of each dNTP, 1μM of each primer, 1U *Taq* polymerase and approximately 100 ng template DNA in a final volume of 12.5μl. (10x Biopro™ buffer contains 160mM $(NH_4)_2SO_4$ and 670mM Tris-HCl (pH 8.8)). Initial denaturation was at 94°C for 2 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 30 seconds. Amplification of *FMN-1* exon 1 and *GREMLIN* exon 1 were particularly difficult therefore Herculase™ Hotstart DNA Polymerase (Stratagene) was preferred. This enzyme consists of *Pfu* DNA polymerase combined with *Taq* polymerase and it provides amplification of particularly difficult GC rich DNA, avoiding the inconvenience of non-specific primer annealing. The amplification reaction contained 1x Herculase polymerase reaction buffer, 0.25mM of each dNTP, 1μM of each primer, 2% DMSO, 1.25U Herculase™ hotstart polymerase and approximately 100ng template DNA in a total volume of 12.5μl. Initial denaturation was at 98°C for 3 minutes, followed by 10 cycles of denaturation at 98°C for 40 seconds, annealing at 58°C for 30 seconds and extension at 72°C for 45 seconds, and a further 25 cycles of denaturation at 98°C for 40 seconds, annealing at 58°C for 30 seconds, extension at 72°C for 55 seconds.

PCR amplification of *SHOX2, Glylu, ZFTF, WISP1* were carried out using HotStarTaq DNA Polymerase (QIAGEN Ltd.), except for *SHOX2* exon 1 amplified with Herculase™ Hotstart DNA Polymerase (Stratagene) as described above (annealing temperature 62°C). HotStarTaq DNA Polymerase is a form of QIAGEN Taq DNA Polymerase modified to provide high specificity in hot-start PCR minimizing non-specific amplification products, primer-dimers and background. Amplification reactions contained 1x PCR buffer, 0.25mM of each dNTP, 1μM of each primer, 2.5U HotStarTaq DNA Polymerase and approximately 100ng template DNA in a total volume of 12.5μl. (10x PCR buffer contains 15mM $MgCl_2$). The initial heat activation step at 95°C for 15 minutes was followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 60-62°C for 30 seconds, extension at 72°C for 30 seconds and a

final step at 72°C for 10 minutes. All PCR reactions were performed in single 0.2ml micro-tubes or 0.2ml strip tubes using a T3 Biometra Thermocycler.

2.3.2.6 DNA amplification using exonuclease III

For amplification of damaged templates like tissue extracted DNA from old paraffin sections, *E.coli* exonuclease III (New England Biolabs) was used according to the manufacturer's instructions. This bacterial exonuclease improves long PCR amplification with DNA samples that are mildly or moderately damaged prior to extraction or during storage. 5, 25 or 50U of exonuclease III was added directly to the PCR mixture along with the DNA polymerases (Fromenty *et al.*, 2000).

2.3.2.7 PCR optimisation and contamination

In order to determine the optimal PCR conditions for new primer pairs, PCR reactions were set up using control DNA. When no PCR product or only a weak PCR product was obtained, or there were non-specific products, the annealing temperature was varied to get optimal amplification. When preparing PCR reactions, standard precautions were taken to avoid contamination of samples. Aerosol resistant ART pipette tips were used and a work area for preparing PCR reagents was set and kept constantly clean. A negative control containing sterile distilled water in place of DNA template was included in most experiments to check for contamination. To assess the success of each PCR, an aliquot of each completed PCR reaction was electrophoresed on a 2% agarose gel.

2.3.3 Agarose gel electrophoresis

PCR products were routinely separated on a 2% NuSieve GTG agarose gel (2g agarose, 100ml 1x TBE buffer, 4µl 10mg/ml ethidium bromide) using a 50 bp DNA ladder as a size standard. Before loading, 2µl orange G loading buffer and 5µl TE buffer were added to 5µl PCR product. Gels were run in 1x TBE buffer containing 4µl 10mg/ml

ethidium bromide per 100ml, at 50-100V for 30-60 minutes depending on the size of fragment and degree of resolution required. DNA fragments were visualised by ethidium bromide fluorescence under UV illumination and photographed. In order to optimise resolution of fragments, higher percentage gels containing 4% NuSieve GTG agarose (4g agarose, 100ml 1x TBE buffer, 4µl 10mg/ml ethidium bromide) were used to separate digested PCR products.

2.3.4 Restriction digestion

Restriction enzyme digestion of two PCR products (ZNF406 exon 3 and exon 4) was necessary to establish if single nucleotide changes found in the sequence were pathogenic mutations by looking at how they segregated in the family, as explained in Chapter 4 (section 4.4.2.9). *HpaII* and *AciI* restriction digestions were therefore carried out according to the manufacturer's recommendations, at the recommended temperature in the minimum volume permitted by the concentration of the DNA using the appropriate dilution of the reaction buffer. Both digestions of PCR products were carried out for two hours. After incubation, DNA fragments were separated on a 4% agarose gel.

2.3.5 Purification of PCR products

2.3.5.1 ExoSAP-IT

For purification of sequencing samples, 5µl of PCR product were purified directly with 2µl of ExoSAP-IT (USB Corporation), in accordance with the manufacturer's instructions. ExoSAP-IT utilizes two hydrolytic enzymes, Exonuclease I and Shrimp Alkaline Phosphatase to remove unwanted dNTPs and primers. The Exonuclease I degrades residual single-stranded primers and any extraneous single-stranded DNA produced by PCR. The Shrimp Alkaline Phosphatase hydrolyses remaining dNTPs from the PCR mixture, which would interfere with the sequencing reaction. ExoSAP-IT

is added directly to the PCR product, incubated at 37°C for 15 minutes for treatment and then at 80°C for further 15 minutes to inactivate the enzymes.

2.3.6 Preparation of gel for linkage and sequence analysis

For electrophoresis on the ABI PRISM™ 377 DNA Sequencer, 4.5% acrylamide gels were prepared. A pair of glass plates (36 cm well-to-read distance) were cleaned with Alconox detergent, washed in distilled water and allowed to air dry upright before being clamped into position in the gel cassette, separated by 0.2mm thick spacers on either side. The gel mix was prepared by adding 18g urea and 0.3g amberlite deionising resin to 27.2ml purified water and 5.6ml 40% acrylamide/bisacrylamide. After mixing by stirring for 10-15 minutes, the resin was removed by vacuum filtration. 5ml filtered 10x TBE buffer was then added to yield 50ml gel mix. To catalyse the polymerisation reaction, 250µl freshly prepared 10% w/v ammonium persulfate and 35µl TEMED were added to the gel mix, prior to pouring the gel between the assembled plates using a 50ml syringe. The flat edge of a plastic casting comb was inserted at the upper end of the gel and clamped into place. The gel was then left to polymerise for at least an hour. After this period, the casting comb was carefully removed and any remaining spots of acrylamide were cleaned from the plates with purified distilled water and 1x TBE buffer. 24, 36 and 48 well disposable shark-tooth combs were used and inserted with the teeth just penetrating the top of the gel. To prevent leakage of samples from one well to another, paper combs were preferred to plastic ones as they swell when in contact with water and fill the gap between wells. The gel cassette was then clamped into the electrophoresis chamber of the automated sequencer, and the upper and lower buffer chambers were filled with 1x TBE buffer.

2.3.7 Linkage analysis

Microsatellite repeat analysis was performed using either the ABI PRISM™ 377 DNA Sequencer or the MegaBACE 1000 DNA Sequencer. Preparation of sample protocols and analysis differ.

2.3.7.1 *Genotyping using the ABI PRISM™ 377 DNA Sequencer*

Fluorescent microsatellite repeats amplified as described in section 2.3.2.4, were electrophoresed and analysed on the ABI PRISM™ 377 DNA Sequencer in accordance with the manufacturer's instructions, using ABI PRISM™ GeneScan® 3.1.2 Analysis software (Perkin Elmer Applied Biosystems Division). 1.5µl diluted PCR product(s) was added to 3.5µl of a loading mix containing GENESCAN-500™ TAMRA, a dye-labelled internal lane standard (displayed as red) which allows extremely accurate sizing of DNA fragments in the 35-500 bp range. The 5µl samples were then denatured at 95°C for 4 minutes and held on ice before loading (1.8µl/well) onto an acrylamide gel pre-heated to 48°C by running for 1 hour at 1,000V. They were electrophoresed for 2-2.5 hours at 3,000V and a gel temperature of 48°C. Results were then analysed using ABI PRISM™ Genescan® 3.1.2 and Genotyper® 2.5 software (Perkin Elmer Applied Biosystems Division) run on an Apple Power Macintosh computer.

2.3.7.2 *Genotyping using MegaBACE™ 1000 DNA Sequencer*

In the last stage of the project analysis of fluorescent microsatellite repeats was performed on the MegaBACE™ 1000 DNA Sequencer (Amersham Biosciences), a high-throughput fluorescence-based DNA analysis system utilizing capillary electrophoresis with 96 capillaries operating simultaneously. PCR products were adequately diluted and 2µl of the diluted product was added to the loading mix containing 0.5µl MegaBACE™ ET400-R (or MegaBACE™ ET550-R depending on size of products) size standard and 7.5µl sterile purified water into a polycarbonate 96 well sample plate. MegaBACE™ ET size standards consists of 20 (ET400-R) and 22 (ET550-R) double stranded DNA fragments in which one strand is end-labelled with an energy transfer (ET) dye. They allow precision sizing of DNA fragments up to 400 or 550 bp respectively. The 10µl samples were then denatured at 95°C for 2 minutes and injected into the MegaBACE sequencer following the manufacturer's instructions. 3 matrix tubes were taken out from storage (4°C) at least an hour before use and

centrifuged at 4000rpm for 4 minutes. A 96 well buffer plate was prepared by adding 200µl 1x linear polyacrylamide (LPA) buffer (Amersham Biosciences) to each well. Prior to prerunning, capillaries were first rinsed with sterile purified water and then filled with matrix buffer. After rinsing the capillaries again with sterile water, the 96-sample plate containing the genotyping products was loaded into the cassette and the samples were injected at 3000V for 45 seconds. The sample plate was then replaced with the buffer plate and the run was started at 10 000V for 75 minutes. Results were analysed with the aid of MegaBACE™ Genetic Profiler software (Amersham Biosciences) which allows automated analysis of di- and tetranucleotide repeats in a Windows™ NT format.

2.3.7.3 Lod score calculations

Two-point lod scores were calculated using the FASTLINK and MLINK programs of the LINKAGE package available at the HGMP website (http://www.hgmp.mrc.ac.uk/). The disease parameters used in the data files were: fully penetrant autosomal recessive with a gene frequency of 0.001. Equal frequencies for each allele identified were used, because population allele frequencies were not available.

2.3.8 Sequencing analysis

2.3.8.1 Dye terminator cycle sequencing reaction

Sequencing reactions were performed using the ABI PRISM Big Dye™ Terminator v2.0 Cycle Sequencing Ready Reaction kit (Perkin Elmer Applied Biosystems Division). In this system instead of radiolabelled ddNTPs, four different, fluorescent dye-labelled chain terminators, recognized as different colours by the automated fluorescence DNA sequencer, are used to terminate DNA synthesis catalysed by AmpliTaq® DNA polymerase. The advantages of this method are that no radioactivity is involved and all four reactions can be performed in one tube and loaded in one gel lane. Sequencing reactions contained 4µl sequencing reagent pre-mix, 5µM sequencing

primer, approximately 50ng PCR product and 1µl of Pellet Paint™ NF Co-Precipitant (Novagen) made up to 10µl with sterile purified water. The primers originally used to amplify the product were employed as forward and reverse sequencing primers. The Pellet Paint™ NF Co-Precipitant is a visible, non-fluorescent, dye-labelled carrier for use in ethanol precipitation of nucleic acids to improve the efficiency and reliability of the precipitation. An initial denaturation step at 96°C for 1 minute was followed by 30 cycles of denaturation at 96°C for 30 seconds, annealing at 50°C for 15 seconds and extension at 60°C for 4 minutes.

2.3.8.2 Ethanol precipitation

Unincorporated ddNTPs were removed from the completed sequencing reactions by ethanol precipitation. 40µl 75% ethanol was first added to each sample. After vortexing, samples were centrifuged at 14,000 rpm for 3 minutes. As much supernatant as possible was aspirated, 125µl 75% ethanol was added to wash the DNA pellet, and the samples were centrifuged at 14,000 rpm for 3 minutes. As much supernatant as possible was aspirated and the samples were left to dry at 37°C for 20 minutes.

2.3.8.3 Sequencing

Samples were electrophoresed and analysed on the ABI PRISM™ 377 DNA Sequencer in accordance with the manufacturer's instructions, using DNA ABI PRISM™ Sequencing Data Collection and Analysis software, run on an Apple Power Macintosh computer. The gel was prepared as described in section 2.3.6 and pre-heated to 51°C by running at 1,000V for an hour. Samples were resuspended in 2µl blue dextran loading buffer, vortexed, spun, denatured at 90°C for 2 minutes and held on ice before loading (1.8µl/well). They were then electrophoresed for 7 hours at 1,680V, using 1x TBE buffer and a gel temperature of 51°C.

The sequence obtained was analysed using the Sequencher™ 4.1 program (GeneCodes). Sequencing files in ABD format along with the known published

sequence were imported into a Sequencher file. The program aligns similar sequences and highlight discrepancies of bases facilitating the finding of mutations and SNPs.

2.3.9 Cloning of PCR products

PCR products were cloned using the PCR-Script™ Amp Cloning kit (Stratagene). This kit was chosen because it permits the efficient cloning of blunt-ended PCR products with a high yield and a low false positive rate. The pPCR-Script Amp SK(+) cloning vector is derived from the pBluescript II SK(+) phagemid. It includes an ampicillin-resistance gene for antibiotic selection of the vector, as well as an *Srf*I restriction-endonuclease target sequence flanked by sites for the M13 (-20) and M13 reverse sequencing primers. The vector is supplied pre-digested with *Srf*I. It also provides a blue-white colour selection of recombinant plasmids since the pPCR-Script cloning vector has an *E. coli lacZ* gene fragment situated in the multiple cloning site, which is disrupted by the presence of an insert. XL10-Gold® Kan ultracompetent cells are a *lac*$^-$ strain with a *lacZΔM15* gene, which is complemented by the *lacZ* gene fragment in the pPCR-Script vector. The *lacZ* gene encodes the enzyme β-galactosidase, which cleaves the substrate X-gal (5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside) when its expression is induced by IPTG (isopropyl-1-thio-β-D-galactopyranoside), resulting in a blue colour. Thus colonies containing a vector without an insert are blue, whereas those with an insert remain white.

The production of high-fidelity, blunt-ended PCR products is ensured by the use of *Pfu* DNA polymerase. HotStarTaq DNA Polymerase was used instead to generate the inserts. To create the blunt ends needed to improve cloning efficiency, purified PCR products were subsequently polished, as recommended by the manufacturer, adding 1μl 10mM dNTP mix (2.5mM each), 1.3μl polishing buffer, 1μl cloned *Pfu* DNA polymerase to 10μl purified PCR product in a screw-topped microcentrifuge tube. After mixing, the polishing reaction was incubated for 30 minutes at 72°C.

For the ligation reaction, 1µl pPCR-Script vector (10ng/µl), 1µl pPCR-Script 10x reaction buffer, 0.5µl 10mM rATP, an appropriate quantity of polished PCR product in 5.5µl elution buffer or water, 1µl *Srf*I (5U/µl) and 1µl T4 DNA ligase (4U/µl) were added in that order to a screw-topped microcentrifuge tube, giving a total volume of 10µl. (*Srf*I is included in the ligation reaction to digest any vector ligated without an insert, maintaining a high concentration of digested vector and thus increasing ligation efficiency.) The quantity of PCR product was calculated to give an insert:vector molar ratio of 40:1 to 100:1, as recommended by the manufacturer, and was 50-150ng, depending on the length of the PCR product. After mixing, the ligation reaction was incubated for 1 hour at room temperature, heated for 10 minutes at 65°C to inactivate the *Srf*I and T4 ligase, and held on ice.

For the transformation, one 200µl vial of XL10-Gold® Kan ultracompetent cells (supplied with the kit and stored at -70°C) were thawed on ice, mixed gently, and placed in 40µl aliquots in pre-chilled 15ml Falcon 2059 polypropylene tubes. 1.6µl XL10-Gold β-mercaptoethanol was added to each tube, and the cell aliquots were incubated on ice for 10 minutes, swirling every 2 minutes. 2µl ligation reaction was then added to each tube, and the cells were swirled gently, incubated on ice for 30 minutes, heat-shocked in a 42°C water bath for 30 seconds, and incubated on ice for a further 2 minutes. Following the addition of 450µl of preheated NZY^+ medium, the cells were incubated for 1 hour at 37°C with shaking at 225 rpm.

About 15 minutes prior to plating the transformations, LB-ampicillin agar plates were spread evenly with 40µl X-gal (25 mg/ml in dimethyl formamide) and 40µl IPTG (25 mg/ml, filter sterilised). Between 50µl and 200µl of the appropriate transformation was then spread on each plate, and the plates were incubated at 37°C overnight. After 1 hour incubation at 4°C next morning to enhance the blue colour, pure white colonies, expected to contain only vector with insert, were selected.

2.3.10 Colony PCR

Colony PCR was used as a method for screening colonies to identify those containing the desired recombinant plasmid. Individual white colonies were picked with a sterile loop and suspended in 30 µl of LB-ampicillin broth. 5 µl of this suspension was removed, incubated at 100 °C for 5 minutes to lyse the cells then diluted in a 10x volume of sterile purified water, This was used as the template for a PCR amplification using a primer pair specific for the insert. For colonies shown to contain the insert, the remaining 25 µl of suspension was inoculated into a 10ml aliquot of LB-ampicillin broth for overnight culture at 37°C at 225 rpm.

For long term storage, a portion of the culture was frozen in the presence of glycerol, to protect cells from ice crystal formation. Glycerol stocks were made by adding 500µl overnight culture to 500µl sterile 50% v/v glycerol in a cryotube, which was then kept at -70°C.

2.3.11 Bioinformatics

For a complete list of URLs used see section 2.3.11.6.

2.3.11.1 Locating microsatellite markers within a region

Polymorphic markers used to establish a dense cover in the genome and markers located in candidate regions were mainly identified using the NCBI MapView at http://www.ncbi.nlm.nih.gov/mapview and the UCSC Human Genome Browser at http://www.genome.ucsc.edu. Both sites provide information on regions of the human genome including genes, clones, ESTs, contigs, genomic markers, STS markers, PCR conditions. They also supply physical, genetic, cytogenetic and radiation maps for map analysis and data integration and sequence variations within the human genome including mutations, polymorphisms, allele frequency and heterozygosity. In the initial

stages of the project linkage map data from The Marshfield Centre for Medical Genetics was primarily used to locate markers, their primer sequences and heterozygosities. Later, the deCode Icelandic map available on the NCBI and UCSC websites was preferred to choose markers. The improved resolution of the deCode map guarantees more accuracy in the genetic distances of markers and provides more highly polymorphic markers, important for the statistical power of linkage analysis. Details of primer sequence, alternative names, heterozygosity, product size and sequence were all available on the NCBI database uniSTS at http://www.ncbi.nlm.nih.gov/genome/sts/.

2.3.11.2 Locating genes within a region

Genes within or near the linked regions were identified using a number of websites including the NCBI MapViewer and the UCSC Genome Browser. They both provide list of known genes lying in a particular region, their DNA and mRNA sequence, BAC clones from which the sequence was derived, protein sequence and function, if known. The UCSC site also gives a list of predicted genes, including Acembly genes, Ensembl genes, RefSeq genes, and also provides links to the NCBI Entrez documents for each clone. Link to OMIM, PubMed and Mouse Genome Informatics are also available. The Ensembl Genome Browser http://www.ensembl.org/ provides more information about gene structure, genomic sequence and intron/exon boundaries, protein sequence and SNPs. More recently easy access to sequence data from other species including mouse, fugu, zebrafish, rats, chimp, mosquito, fruitfly have also been made available.

2.3.11.3 Database homology searches

In order to identify the genomic sequence of a particular gene when only the mRNA sequence was known, BLAST (Basic Local Alignment Search Tool) searches located either at http://www.ncbi.nlm.nih.gov/BLAST/ or at http://www.ensembl.org/ and BLAT searches available on the UCSC website were performed against the human genomic sequence using the FASTA formatted mRNA sequence obtained from Entrez. BLAST searches were also used to identify homologous genes to candidate genes in the genome.

2.3.11.4 SNPs database

The Database of Single Nucleotide Polymorphisms (dbSNP) located at http://www.ncbi.nlm.nih.gov/SNP was used to verify if changes in the genomic sequence identified during sequencing of candidate genes were known SNPs or not.

2.3.11.5 Protein prediction program

In order to determine the structure of a protein from a given nucleotide sequence, the sequence was translated using the SIXFRAME translator program available at http://workbench.sdsc.edu and the structure of the protein determined using the PROSITE (http://www.expasy.org/prosite/) and PSORT (http://psort.nibb.ac.jp/) protein prediction programs.

2.3.11.6 URLs

PRIMER3-Design primers pairs and probes

http://workbench.sdsc.edu

Databases homology searches

http://www.ncbi.nlm.nih.gov/BLAST

NCBI

http://www.ncbi.nlm.nih.gov/

University of California, Santa Cruz, Genome Browser Gateway

http://genome.cse.ucsc.edu/

Ensembl human genome server

http://www.ensembl.org/

OMIM

http://www.ncbi.nlm.nih.gov/Omim/

Entrez-PubMed

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi

Entrez nucleotide search

http://www.ncbi.nlm.nih.gov/entrez/nucleotide.html

UK HGMP-RC

http://www.hgmp.mrc.ac.uk/

The Genome Database (GBD)

http://gdbwww.gdb.org/

Whitehead Institute

http://www-genome.wi.mit.edu

Unigene

http://www.ncbi.nlm.nih.gov/UniGene/index.html

Human/mouse homology

http://www.ncbi.nlm.nih.gov/Omim/Homology

SDSC Biology Workbench

http://workbench.sdsc.edu

Restriction sites

http://workbench.sdsc.edu

UniSTS

http://www.ncbi.nlm.nih.gov/genome/sts/

Map View

http://www.ncbi.nlm.nih.gov/mapview

PROSITE

http://www.expasy.org/prosite/

PSORT

http://psort.nibb.ac.jp/

Chapter 3 Cenani-Lenz syndrome

The limb abnormalities in CLS patients were first noted to be very similar to those found in the recessive *ld* mouse mutant by Professor Robin Winter (Winter, 1988). From what was known about *ld* mutants at the commencement of this project (see section 1.2.3.1), human *FORMIN-1* (*FMN-1*) made a promising candidate gene for CLS. Following recent findings showing that the *ld* phenotype is caused by disruption of regulatory elements controlling *Gremlin* transcription in the limb bud mesenchyme or by direct disruption of the *Gremlin* transcription unit (section 1.2.3.2), human *GREMLIN* became a second excellent candidate gene for CLS.

3.1 Results

3.1.1 Haplotype analysis

FMN-1 lies in a 1.9 cM interval on chromosome 15q13.2 between markers D15S1010 and D15S144 (between 30.89-30.97 Mb)(Richard *et al.*, 1992). Interestingly human *GREMLIN* lies in the same chromosomal region (15q13-q15), extremely close to *FMN-1* at 30.72-30.74 Mb (Topol *et al.*, 2000). In order to confirm whether the region containing *FMN-1* and *GREMLIN* was linked to the CLS locus, haplotype analysis was performed in consanguineous Families A, B, C and D. Nine polymorphic markers were chosen to cover about 12 cM in the region of interest. Two were taken from the linkage sets available in the laboratory (D15S165 from the ABI set and ACTC from the Research Genetics set) (see section 2.3.2.1), while primers to a further seven markers were chosen from the NCBI UniSTS database (see section 2.3.11.1) and designed as explained in section 2.3.2.2. Markers were chosen based on map position and degree of heterozygosity. These were D15S976, D15S1031, D15S1010, D15S144, D15S1007, D15S1040 and D15S118. Details of these markers are shown in

Table 2.1 and their positions in Figure 3.1. The centimorgan distances quoted for the Marshfield genetics maps were used as they were available for all markers. The physical distances shown in Figure 3.1 correspond to the more recent NCBI genome assembly build 34.

PCR to amplify microsatellite repeats was carried out as described in section 2.3.2.4 and genotyping analysis as described in section 2.3.7.1.

3.1.1.1 Family A

Family A was the first family from whom DNA was obtained. Initial haplotype analysis across the chromosome 15q region where *FMN-1* and *GREMLIN* lie, showed that patient II.2 was heterozygous for all nine markers, thereby excluding a recessively-acting inherited-by-descent (IBD) mutation in either *FMN-1* or *GREMLIN* as the cause of her limb and renal abnormalities (Fig. 3.2). The results were reported in (Bacchelli *et al.*, 2001).

3.1.1.2 Family B

DNA from Family B was obtained in a second instance. Marker D15S976 was found to be located on chromosome 3 and not on chromosome 15 following revision of the human genomic sequence and therefore it was not included in the haplotype analysis of Families B, C and D. Results in Family B showed homozygosity across the region in all eight markers tested (Fig. 3.3). The affected child II.4 was homozygous across the region while his parents (I.1 and I.2) and three siblings (II.1, II.2 and II.3) were heterozygous, suggesting that the affected child might carry an IBD mutation in either *FMN-1* or *GREMLIN.*

3.1.1.3 Family C and Family D

Haplotype analysis in Families C and D was performed using four polymorphic markers in the chromosome 15q region (Fig. 3.4). Both probands were heterozygous

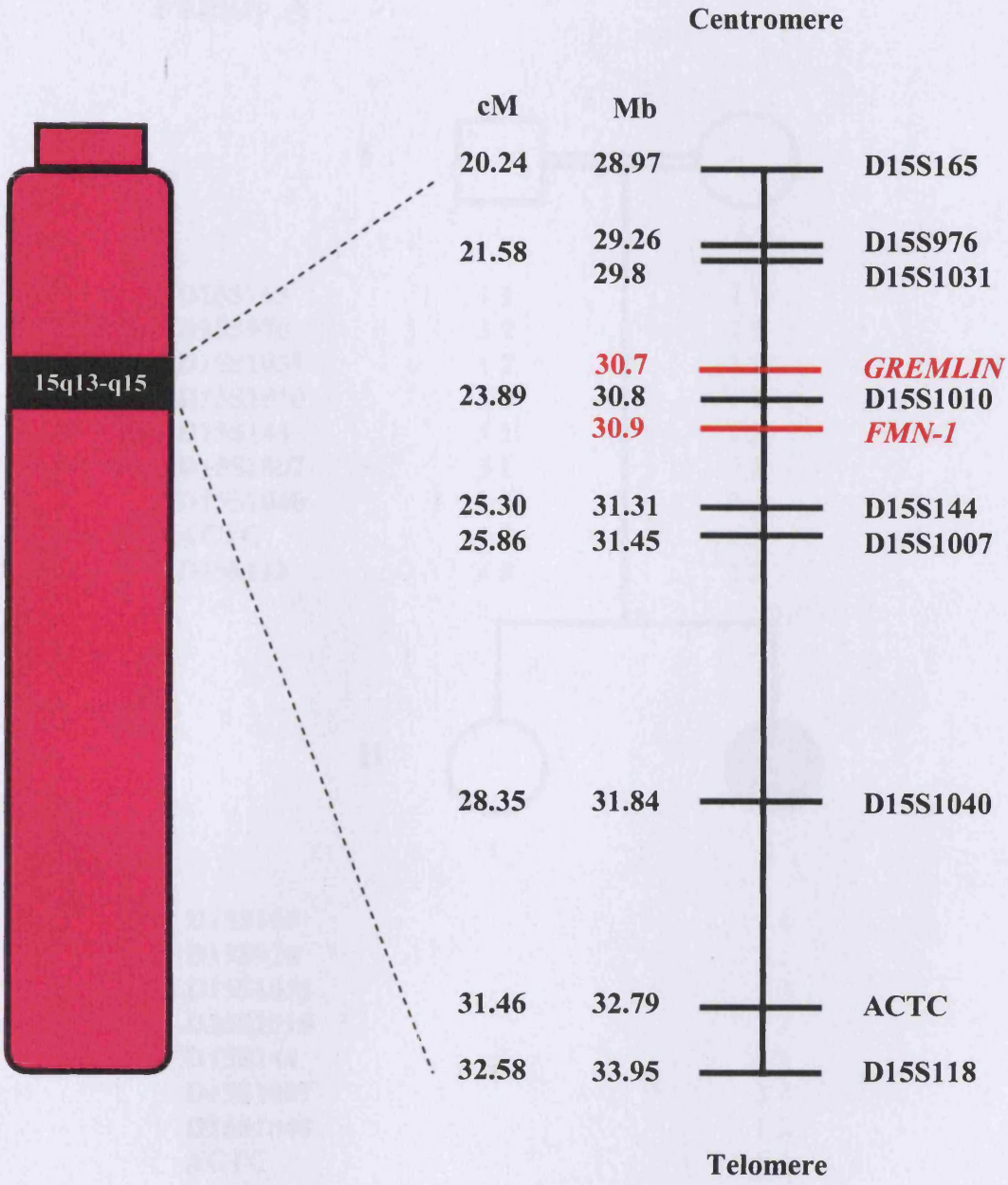
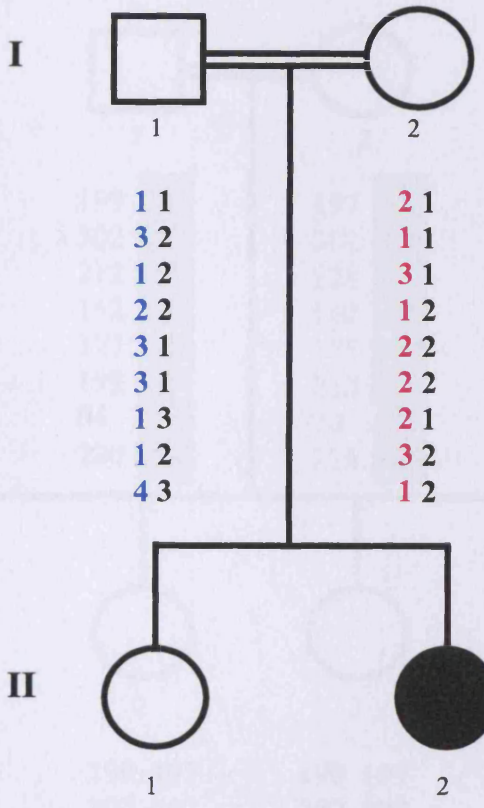


Figure 3.1 Schematic representation of the region on chromosome 15q13-q15 where *FMN-1* and *GREMLIN* lie. Genetic and physical positions of polymorphic markers used for haplotype analysis in consanguineous Families A, B, C and D are given.

Family A



	I.1	I.2
D15S165	1 1	2 1
D15S976	3 2	1 1
D15S1031	1 2	3 1
D15S1010	2 2	1 2
D15S144	3 1	2 2
D15S1007	3 1	2 2
D15S1040	1 3	2 1
ACTC	1 2	3 2
D15S118	4 3	1 2

	II.2
D15S165	1 2
D15S976	3 1
D15S1031	1 3
D15S1010	2 1
D15S144	3 2
D15S1007	3 2
D15S1040	1 2
ACTC	1 3
D15S118	4 1

Figure 3.2 Results of haplotype analysis in Family A using 9 polymorphic markers on chromosome 15q13-q15. Father (I.1) was informative and heterozygous for all the markers except D15S165 and D15S1010; the mother (I.2) was informative for all the markers except D15S976, D15S144 and D15S1007.

Family B

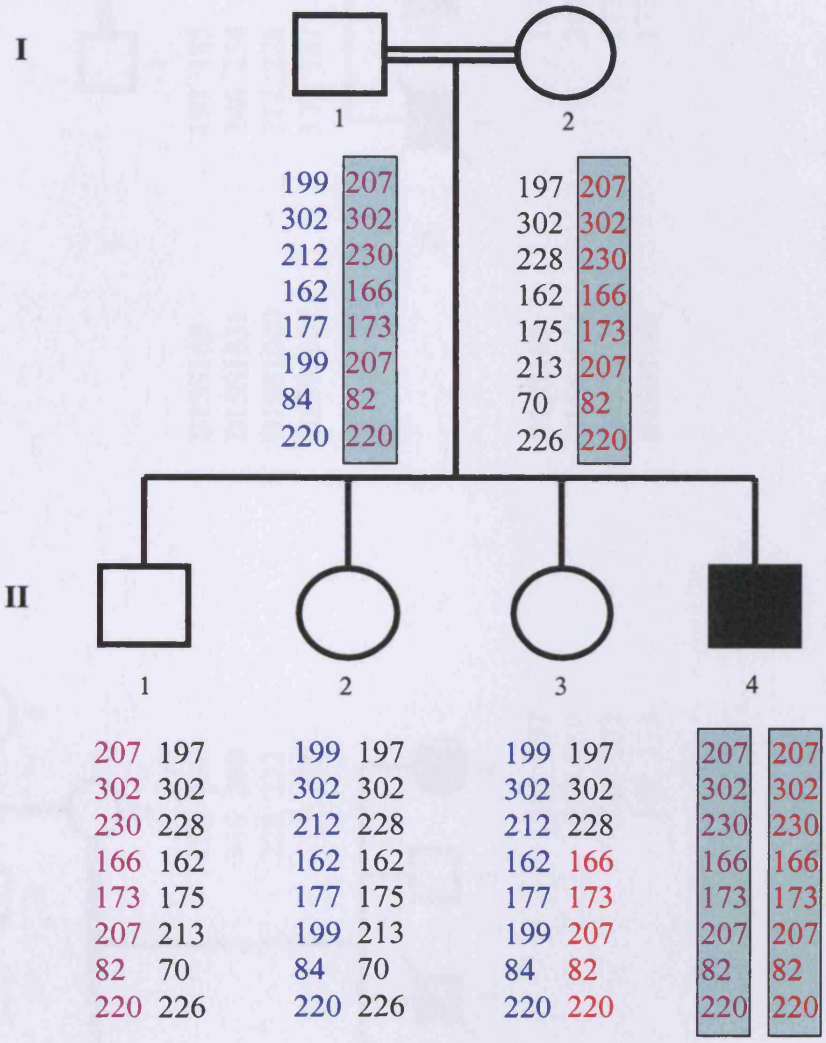
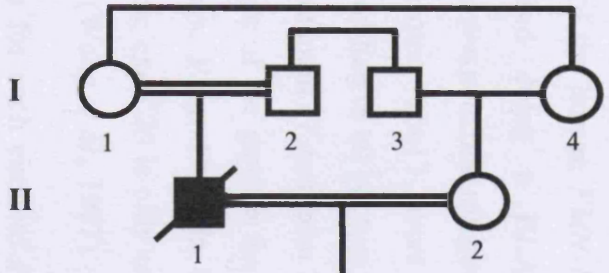
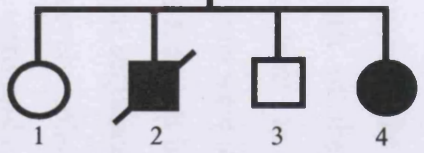


Figure 3.3 Results of haplotype analysis in the 15q13-q15 region in Family B. Note homozygous region in patient II.4 between D15S165 and D15S118. Both parents and the three unaffected siblings were heterozygous across the region, except for marker D15S1031 which was non-informative.

Family C

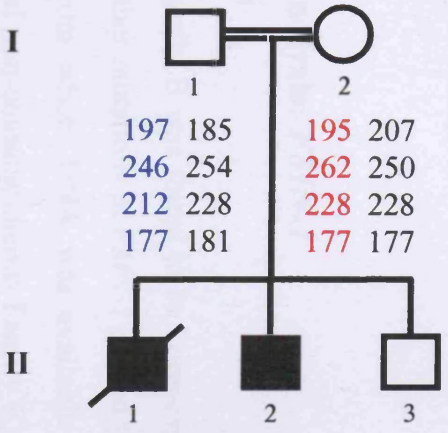


D15S165	183	197
D15S1031	250	250
D15S1010	228	232
D15S1007	173	173



D15S165	183	197
D15S1031	260	250
D15S1010	228	232
D15S1007	169	173

Family D



D15S165	197	185	195	207
D15S1031	246	254	262	250
D15S1010	212	228	228	228
D15S1007	177	181	177	177

D15S165	197	195
D15S1031	246	262
D15S1010	212	228
D15S1007	177	177

Figure 3.4 Results of haplotype analysis in the 15q region in Families C and D. Both probands (III.4 in Family C and II.2 in Family D) were heterozygous across the region.

across the region, thus excluding a recessively-acting IBD mutation in either *FMN-1* or *GREMLIN* in these two families.

3.1.2 Sequence analysis: *FMN-1*

The proband from Family B was found to be homozygous across the 15q region. In order to identify whether mutations in *FMN-1* caused CLS, *FMN-1* was sequenced in the three families from whom DNA was available at the time (consanguineous Families A and B and non-consanguineous Family E). DNA from Families C, D, F and G was obtained later and kept in storage.

3.1.2.1 Identification of exons in human FMN-1

The mRNA sequence of mouse *Formin-1* was available in the databases (GenBank Accession No X53599) and the intron/exon boundaries had already been identified (Wang *et al.*, 1997), but the human *FMN-1* sequence was not available. Therefore, exons of the human *FMN-1* gene were identified by taking each individual mouse exon and doing a BLAST search against the human genomic sequence (http://www.ncbi.nlm.nih.gov/BLAST/). Most of the relevant regions on human chromosome 15q13 were available, at least in draft form. Human exons corresponding to all but two of the mouse exons were identified. Exons corresponding to mouse exon 17 and exon 20 could not be found. It is possible that exons 17 and 20 lie in one of the gaps in the available human sequence or they may not have human homologs. Exon 17 is in fact an alternatively spliced exon containing repetitive sequence; exon 20 is only expressed in the testis and has stop codons in all 3 reading frames (Wang *et al.*, 1997).

Primers for each exon of *FMN-1* were designed and synthethised as described in section 2.3.2.3. Primer details and PCR conditions are shown in Table 2.9. Primers were spaced so that splice sites as well as exons could be seen. Overlapping primers were designed to cover large exons. The last coding exon, exon 24, was examined up

to and just beyond the STOP codon. Most of the 3'UTR, which is approximately 5 kb, was not sequenced.

3.1.2.2 PCR amplification of FMN-1

PCR amplification of *FMN-1* was carried out as described in section 2.3.2.5. Each exon was successfully amplified by PCR in both the carrier father (I.1) and affected child (II.4) from Family B as well as in patients II.2 from Family A and II.1 from Family E. A single product of the expected size was obtained for each individual. Therefore, no evidence of a small insertion or deletion within any of the exons was detected on PCR in any of the patients and no single exon was homozygously deleted in child II.4 from Family B.

3.1.2.3 FMN-1 sequencing results

All twenty-two exons were mutation screened by direct sequencing using protocols detailed in section 2.3.8 and analysed using Sequencher 4.1 (see section 2.3.8.3).

(i) Family B

FMN-1 was first sequenced in Family B in the carrier father (I.1), as it is easier to see any sequence alteration in a heterozygote than in a homozygote. On sequence analysis, he was found to be heterozygous at twelve different positions scattered throughout the gene (Fig. 3.5) confirming that two alleles were present across most of the gene. Two of these were poly-T repeats, one upstream of exon 11 and one upstream of exon 13. A further ten positions at which he was heterozygous were identified. Four of these were in non-coding regions. Six were in coding regions. Three of these do not change an amino acid but three do.

Sequencing of the exons where a change in I.1 had been identified was then performed in the affected child (II.4). The results are shown in Figure 3.5. II.4 was

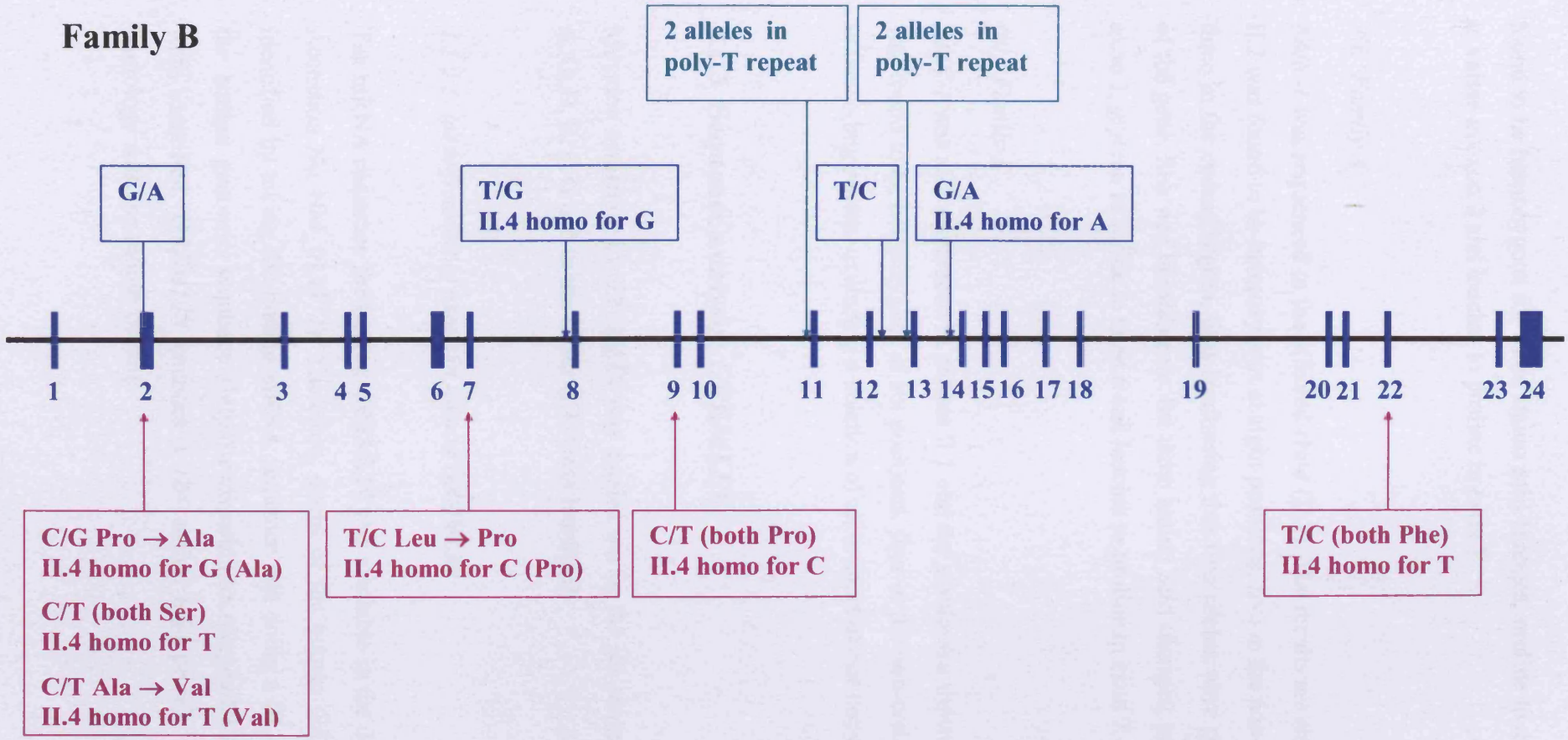


Figure 3.5 Sequence changes in *FMN-1* in I.1 and II.4 from Family B. Changes in non-coding regions are shown in blue and green boxes above the exons. Changes in coding regions are shown in purple boxes below the exons. Inside each box, the result for I.1 is given above the results for II.4. Not all the regions sequenced in I.1 were sequenced in II.4.

found to be homozygous for three amino acid changes, proline to alanine and alanine to valine in exon 2 and leucine to proline in exon 7.

(ii) Family A

FMN-1 was sequenced in the affected child (II.2). The results are shown in Figure 3.6. II.2 was found to be heterozygous at eight positions, five in the non-coding region and three in the coding region, thus confirming that two alleles were present across most of the gene. She was homozygous for three amino acid changes, proline to alanine in exon 2, glycine to valine in exon 6 and leucine to proline in exon 7.

(iii) Family E

FMN-1 was also sequenced in patient II.1 and the results are shown in Figure 3.7. He was found to be heterozygous at six positions, four in the non-coding region and two in the coding region, excluding a deletion of all or most of one copy of the gene.

3.1.3 Sequence analysis: *GREMLIN*

Mutation screening in *GREMLIN* was carried out in the six probands from Families A, C, D, E, F, G and in the father (I.1) from Family B.

3.1.3.1 Identification of exons in human GREMLIN

The mRNA sequence for human *GREMLIN* was available in the databases (GenBank Accession No NM_013372). Therefore, exons of the human *GREMLIN* gene were identified by taking the human mRNA sequence and doing a BLAST search against the human genomic sequence (http://www.ncbi.nlm.nih.gov/BLAST/). Two exons were identified. *GREMLIN* encodes a 184 amino acid protein which shows high homology with the mouse protein.

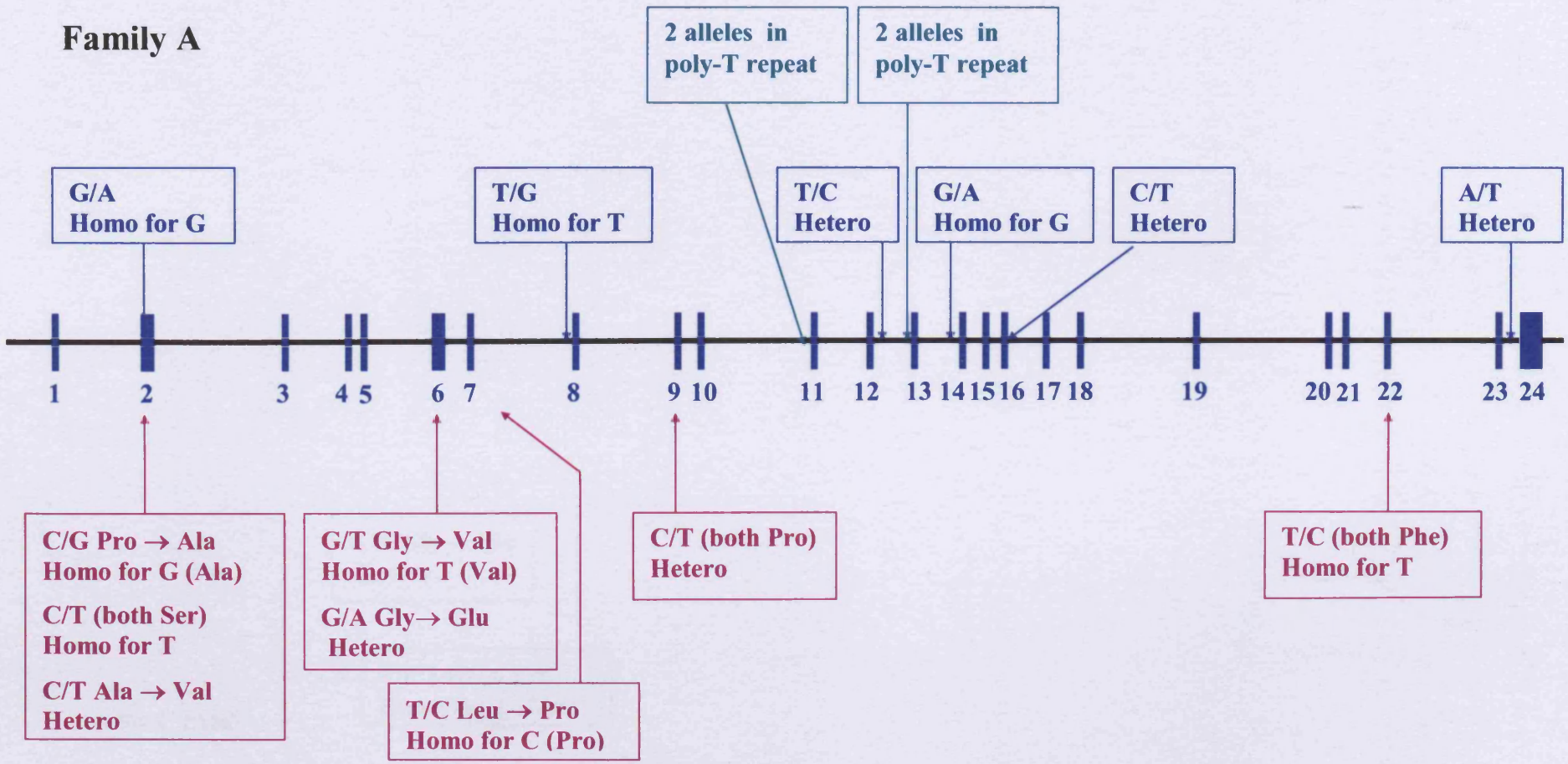


Figure 3.6 Sequence changes in *FMN-1* in II.2 from Family A. Changes in non-coding regions are shown in blue and green boxes above the exons. Changes in coding regions are shown in purple boxes below the exons.

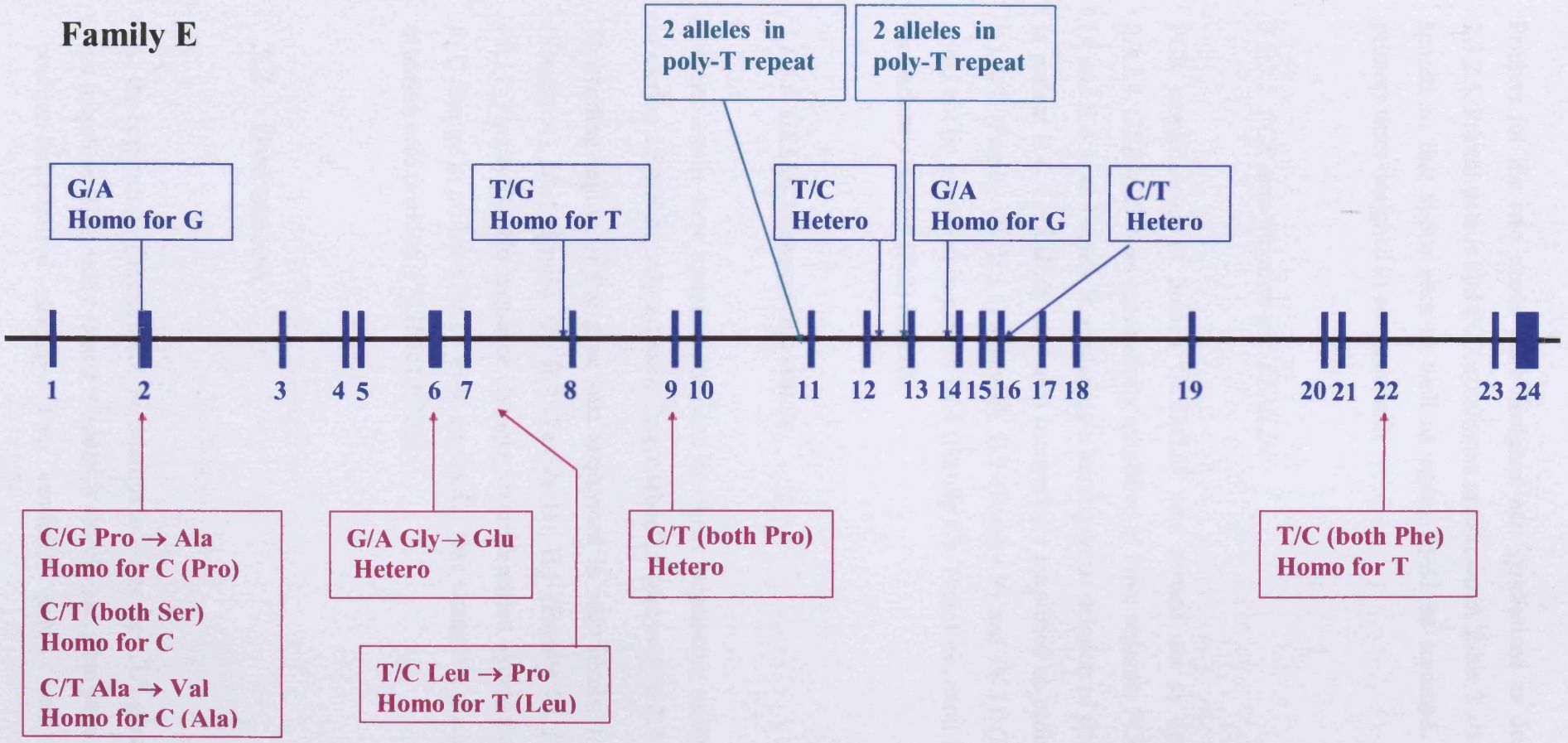


Figure 3.7 Sequence changes in *FMN-1* in II.1 from Family E. Changes in non-coding regions are shown in blue and green boxes above the exons. Changes in coding regions are shown in purple boxes below the exons.

Primers for the two exons were designed and synthesised as described in section 2.3.2.3. Primer details and PCR conditions are shown in Table 2.10. The primers were spaced so that splice sites as well as exons could be screened. Two overlapping primers were designed to cover exon 2.

3.1.3.2 PCR amplification of GREMLIN

PCR amplification of human *GREMLIN* was carried out as described in section 2.3.2.5. *GREMLIN* was successfully amplified in three separate PCR reactions in both I.1 and II.4 from Family B, excluding a homozygous deletion of part or all of the gene in patient II.4. *GREMLIN* was also successfully amplified in probands: II.2 (Family A), II.2 (Family D), II.1 (Family E), II.1 (Family F) and IV.11 (Family G). Exon 1 could not be amplified in patient III.4 (Family C). Therefore, exon 1 was sequenced in the patient's mother (II.2) instead.

3.1.3.3 GREMLIN sequencing results

The two exons were mutation screened by direct sequencing using protocols detailed in section 2.3.8 and analysed using Sequencher 4.1 (see section 2.3.8.3).

The coding region of the gene was sequenced in individuals: I.1 (Family B), II.2 (Family A), III.4 (Family C), II.2 (Family D), II.1 (Family E), II.1 (Family F) and IV.11 (Family G). No sequence changes were identified except for a heterozygous G to C change in proband IV.11 from family G. This changes a glutamine to a histidine at amino acid position 175 (Glu175His).

3.2 Discussion

At the beginning of the project, no attempt to locate the CLS gene had been reported in the literature. This study aimed to identify the disease gene causing CLS by adopting a position-independent strategy. Two candidate genes were identified based on

similarities in the phenotype between CLS patients and a mouse model, the *ld* mouse mutant (section 1.2.3). *Ld* mouse phenotypes were until very recently thought to arise from mutations in the *formin-1* gene, suggesting that human *FMN-1* on chromosome 15q13.2 was a good candidate gene for CLS. *FMN-1* was screened for causative mutations by direct sequencing of coding regions in three CLS patients. The BMP-antagonist *Gremlin* lies very near to *formin-1* in both mouse and human genomes. Human *GREMLIN* was also identified as an excellent candidate gene for CLS following recent findings demonstrating that the disruption of *Gremlin* leads to the *ld* phenotype (section 1.2.3.2). *GREMLIN* was therefore screened for causative mutations by direct sequencing of coding regions in seven CLS patients.

3.2.1 Haplotype analysis

Haplotype analysis in Family A showed that the affected child has inherited different alleles from her first-cousin parents across a 12 cM region containing both *FMN-1* and *GREMLIN*. These results excluded the presence of a recessively-acting IBD mutation in either *FMN-1* or *GREMLIN* in this child. At a later stage of the project, affected children from consanguineous Families C and D were also found to be heterozygous across the *GREMLIN / FMN-1* region.

However, in Family B, the affected child was found to be homozygous by descent across the same 12 cM region containing *FMN-1* and *GREMLIN*, while the three unaffected siblings and the parents were all heterozygous across this region, suggesting that the affected child might carry an IBD mutation in either *FMN-1* or *GREMLIN*.

The lack of homozygosity in the affected children in Families A, C and D could then be explained by locus heterogeneity. Alternatively, they might still carry mutations in both copies of *FMN-1* (*or GREMLIN*), but instead of being two copies of the same mutation, inherited from a common ancestor, these could be two different mutations, one inherited from the father's side, and one inherited from the mother's side (compound heterozygotes). However, this is unlikely to occur in three out of four families.

3.2.2 Mutation screening of *FMN-1*

3.2.2.1 Family B

FMN-1 was therefore amplified by PCR in both the carrier father and the affected child from Family B and the entire coding region of the gene was sequenced in the carrier father. Selected regions of *FMN-1* were then sequenced in the affected child. The PCR results excluded the possibility that the child has a homozygous deletion that removes all of *FMN-1*. On sequence analysis, the father was also heterozygous at twelve different positions in the gene (four in exon 2, one in exon 7, one upstream of exon 8, one in exon 9, one upstream of exon 11, two in the intron between exons 12 and 13, one upstream of exon 14, and one on exon 22). This excluded the possibility that the child has a homozygous deletion that removes any of the following individual exons: 2, 7, 8, 9, 11, 12, 13, 14 and 22. As expected, the child was homozygous at all the positions where the father was heterozygous. Four of these positions lie in non-coding regions, but they are not near enough to intron/exon boundaries to affect a splice site directly. Six of these positions lie in coding regions, but only three of them change an amino acid:

(i) Pro to Ala in exon 2. This could potentially be significant, as it is not a conservative change, and mouse *formin-1* has Pro at this position too.

(ii) Ala to Val in exon 2. This is not likely to be significant, as it is a conservative change. Also, mouse *formin-1* has Thr at this position, and the surrounding region of exon 2 is not strongly conserved between human and mouse.

(iii) Leu to Pro in exon 7. This is also not likely to be significant, even though it is not a conservative change, as mouse *formin-1* has Pro at this position.

3.2.2.2 Families A and E

FMN-1 was then amplified and sequenced in two further CLS patients, the affected children in Families A and E. In both cases, all exons were amplified successfully, so at least one copy of each exon is not deleted. In Family A, the affected child was heterozygous at eight positions (one in exon 2, one in exon 6, one in exon 9, one upstream of exon 11, two in the intron between exons 12 and 13, one downstream of exon 16 and one upstream of exon 24). This excluded the possibility of a deletion affecting any of the following exons: 2, 6, 9, 11, 12, 13, 16 and 24. In Family E, the affected child was heterozygous at six positions (one in exon 6, one in exon 9, one upstream of exon 11, two in the intron between exons 12 and 13 and one downstream of exon 16). This excluded the possibility of a deletion affecting any of the following exons: 6, 9, 11, 12, 13 and 16. Both children had several base changes that affect the amino acid sequence:

(i) Pro to Ala in exon 2. The affected child in Family A is homozygous for Ala, like the child in Family B. The affected child in Family E is homozygous for Pro.

(ii) Ala to Val in exon 2. The affected child in Family A is heterozygous. The affected the child in Family E is homozygous for Ala.

(iii) Leu to Pro in exon 7. The affected child in Family A is homozygous for Pro, like child in Family B. The affected child in Family E is homozygous for Leu.

Also, the affected children in families A and B both are heterozygous for a Gly to Glu change in exon 6, a non-conservative change. Mouse *formin-1* has Gly at this position, but the surrounding region of exon 6 is not strongly conserved between human and mouse. The affected child in family A is also homozygous for a Gly to Val change in exon 6, a conservative change. Mouse *formin-1* has Val at this position. Therefore, the change is probably not significant.

3.2.2.3 Conclusion

The entire coding region of *FMN-1* has been sequenced in three unrelated CLS patients, and there is no obvious mutation, such as a premature stop, an insertion, a deletion, or a splice-site mutation, in any of them. However, one or more of the base changes identified could in principle represent a mutation. One or more of the amino acid changes might be significant, but they all seem more likely to be just polymorphisms. This could be confirmed by investigating how often they occur in a panel of normal controls (by direct sequencing and/or by using a restriction digest). Other possibilities are that one or more of the base changes might create a cryptic splice site or one of the exonic base changes (including the ones that do not change an amino acid) could affect an exonic splicing enhancer or an exonic splicing silencer site, leading to erroneous inclusion or skipping of an exon (Liu *et al.*, 2001). This could be investigated by looking for a splice-site abnormality by RT-PCR, but would require obtaining cell lines from the affected children, and these are not currently available. It is also possible that a mutation is present, but could not be found. In fact, not all of the introns, the 3'UTR or the 5'UTR were sequenced. The promoter regions were not sequenced either, mainly because they were poorly characterised. Finally, chromosomal rearrangement in/near the gene using FISH and/or Southern blots were not looked for, but this would again require obtaining cell lines from the affected children.

The fact that out of four consanguineous families, only the affected child of Family B is homozygous across the *GREMLIN / FMN-1* region, could suggest either locus heterogeneity or that the haplotype result in this family happened by chance. False positive results like this are a known pitfall in homozygosity mapping (Miano *et al.*, 2000).

3.2.3 Gremlin role in limb patterning

The probable role of mouse formin-1 and other formin homology proteins in organizing the cytoskeleton makes it difficult to explain why disruption of the cytoskeleton would

lead to loss of *Gremlin* expression and interruption of *Shh* signalling from the ZPA to the AER. During the course of the project, an important paper on Gremlin was published in *Nature Genetics* (Khokha *et al.*, 2003). It was already known that BMP antagonists such as Noggin, Gremlin, DAN, Chordin and Follistatin had a number of different roles during limb development (Merino *et al.*, 1999; Zuniga *et al.*, 1999). Khokha *et al.* (2003) were the first to make a *Gremlin* KO mouse. The Shh-Fgf feedback loop was disrupted in *Gremlin* mutant mice indicating an important role of Gremlin in maintaining the AER and the Shh-Fgf feedback loop during early limb development and patterning. Interestingly, homozygous *Gremlin* KO mice have an identical phenotype to homozygous *ld* mice which results from mutations within the *formin-1* gene (see section 1.2.3). The limb and kidney abnormalities were exactly the same as in *ld* mice. Furthermore, mice heterozygous for both the *Gremlin* null mutation and the *ld* deformity mutation have the full *ld* phenotype showing that the *Gremlin* null mutation is actually allelic to *ld.*

The relationship between *Gremlin* and *formin-1* is interesting because in mice, as in humans, they lie very close together, only 40 kb apart. Khokha *et al.* (2003) favoured the idea that the mutations previously identified in *ld* mice act not by disrupting *formin-1*, and therefore indirectly affecting *Gremlin* expression as previously thought, but by disrupting cis-regulatory elements for *Gremlin* that lie within the *formin-1* gene, thereby affecting *Gremlin* expression directly. This hypothesis was later confirmed by Zuniga and colleagues (Zuniga *et al.*, 2004). As explained in section 1.2.3.2, the authors identified a shared *cis*-regulatory element within *Formin-1* genomic regions that is required for both *Formin-1* and *Gremlin* expression as well as for *Gremlin* activation in the posterior limb bud mesenchyme (Zuniga *et al.*, 2004)

3.2.4 Mutation screening of *GREMLIN*

After the paper by Khokha *et al.* (2003) was published, *GREMLIN* was thought to be a more plausible candidate gene for CLS than *FMN-1*. DNA samples from

consanguineous Families C and D as well as from non-consanguineous Families F and G was obtained in the meantime.

It was then decided to mutation screen *GREMLIN* in all the probands from the new families as well as in the two probands from Families A and E and the father (I.1) from Family B. At the same time, haplotype analysis in consanguineous Families C and D was carried out using four polymorphic markers in the region and linkage to the *GREMLIN / FMN-1* locus was excluded in both families.

The two coding exons of *Gremlin* were successfully amplified in all patients, thus excluding the possibility of a deletion in the gene. The only exception was patient III.4 (Family C). Exon 1 could not be amplified by PCR in this patient, suggesting the possible deletion of this exon. No mutations were found in I.1 (Family B) and in any of the affected patients, except for a heterozygous G to C change in proband IV.11 from Family G which changes a glutamine to a histidine at amino acid position 175. The SNP Database does not report this as a known SNP, and the residue in question is glutamine in both mouse and fugu. However, comparison of the protein sequence of other BMP antagonists, in the Conserved Domain Database at NCBI, showed that two DAN-domain proteins closely related to Gremlin have a histidine instead of glutamine at this position. Firstly, Prdc (protein related to DAN and Cerberus, gi 3252785) and secondly, the DAN zinc finger protein Nbl1 (gi 2498287) mouse neuroblastoma suppressor of tumorigenicity 1 precursor (Fig 3.8). Five other DAN-domain proteins have glutamic acid at that position. So, the change identified in proband IV.11 from Family G is probably just a non-reported SNP. Patient III.4 (Family C) needs to be investigated further for a possible deletion of exon 1, for example by Southern blot hybridization.

Even though no mutations were identified in the coding region of *GREMLIN* in seven CLS patients, an intronic mutation or a regulatory mutation affecting *GREMLIN*, like those identified in the *ld* mice, cannot be excluded (Zuniga *et al.*, 2004). An example is given by chromosome 7q36 preaxial polydactyly (PPD), a commonly observed human congenital hand malformation (Lettice *et al.*, 2003). Point mutations leading to the PPD phenotype were identified in a *Shh* regulatory element located within intron 5 of the

		H	
Human GREMLIN	172	RVKQCRCISIDLD	184
Mouse Gremlin	172	RVKQCRCISIDLD	184
Fugu Gremlin	160	RVKQCRCISIDLD	173
Mouse Prdc	151	KVKHCRCMSVNLS	163
Mouse Nbl1	113	KIVHCSCQACGKE	125

Figure 3.8 Gremlin amino acids sequence around the Glu175His change found in patient IV.11 (Family G) in human, mouse and fugu. Two mouse DAN-domain proteins, Prdc and Nbl1, have histidine instead of glutamine.

Lmbr1 gene, 1 Mb upstream of *Shh* itself, in both human and a corresponding mouse mutant *hemimelic extra toes* (*Hx*).

3.2.5 CLS phenotype and locus heterogeneity

3.2.5.1 Families C and D

The two Egyptian consanguineous Families C and D described by Temtamy *et al.* (2003) do not show the typical CLS phenotype (see section 2.1.2.1). In both families, the very severe limb abnormalities are associated with facial dysmorphism, dental abnormalities and partial tongue-tie, which have not previously been reported in CLS. The feet are very severely affected in patients from both families, while feet are usually less affected than hands or show no abnormalities at all in typical CLS. This may reflect locus heterogeneity and could be the reason why the two consanguineous families are not linked to the *GREMLIN / FMN-1* locus. Alternatively, these two families could have been mis-diagnosed and could represent a separate condition.

3.2.5.2 Family G

In Family G, the proband (IV.11) also has very severe feet abnormalities and congenital cataract not previously associated with CLS (see section 2.1.2.1)(Percin and Percin, 2003). The proband's mother has hypoplasia of the middle and distal phalanges of the second to fifth toes like the proband, which could suggest a manifestation of hererozygosity for a mutation in the CLS gene.

In the same family, a half-cousin (V.2) of the proband has an unusual form of mesoaxial syndactyly similar to syndactyly type 1 (SD1). Her parents are consanguineous, so it is possible that she has a homozygous form of SD1. However, neither of her parents have any limb abnormalities, although SD1 is known to exhibit incomplete penetrance. The locus for SD1 has been mapped to a 9.4 cM interval on

chromosome 2q34-q36 (Bosse *et al.*, 2000), so it would be useful to carry out linkage analysis with markers from this region in patient V.2, her parents and sibs and look for homozygosity-by-descent in V.2.

Although the authors mention the possibility of intrafamilial variable expressivity, it is more likely that individuals IV.11 and V.2 represent two separate conditions occurring in the same family, as they are phenotypically quite dissimilar. Furthermore, no instances of severe SD1 have previously been reported amongst multiply affected CLS sibships.

3.2.5.3 Families A and F

In Family A, the affected child (II.2) has been described with renal abnormalities very similar to those described in *ld* mice. However, such defects have never previously been reported in CLS (see section 2.1.2.1)(Bacchelli *et al.*, 2001).

The proband (II.1) from Family F also has atypical features not previously described in CLS, such as rib and vertebral abnormalities as well as mixed hearing loss (see section 2.1.2.1)(Seven *et al.*, 2000).

The atypical features in Families A, C, D, F and G widen the CLS phenotypic spectrum. Phenotypic variability is commonly described in human disorders and could reflect differences in the genetic background, or be due to environmental factors. The variability could also be of some significance in CLS, indicating locus heterogeneity and the possibility that mutations in more than one gene lead to the CLS phenotype.

3.2.6 Conclusion

In conclusion, out of four consanguineous CLS families, only one was consistent with linkage to the *GREMLIN / FMN-1* region, but no convincing mutations have been

found in the coding region of the two genes in this family. However, it is possible that one or more of the changes identified in *FMN-1* in this family represent a real mutation, but it is also possible that the haplotype result in this family happened by chance. *FMN-1* has not been mutation screened in all the CLS patients and not all the introns, the 3'UTR or the 5'UTR and the promoter regions were sequenced. Therefore, it is still possible that mutations in *FMN-1* have been missed.

The recent findings on *Gremlin* by Khokha and colleagues (Khokha *et al.*, 2003) and Zuniga and colleagues (Zuniga *et al.*, 2004), made *GREMLIN* a stronger candidate gene for CLS. However, no mutations have been identified in the coding region of human *GREMLIN* in six CLS cases. Patient III.4 (Family C) needs to be investigated further for a possible deletion of exon 1.

Despite the fact that mutations in *GREMLIN* and *FMN-1* coding regions have not been identified in our panel of patients, mutations in the homologous *cis*-regulatory elements identified by Zuniga and colleagues may be the cause of CLS malformation.

3.2.7 Future work

More CLS families need to be ascertained. The best way to find out if *GREMLIN* really does have a role in CLS would be to sequence *GREMLIN* in further CLS patients to see if a convincing causative mutation could be found. Investigation of the homologous *cis*-regulatory elements in *FMN-1* also needs to be carried out.

However, other candidate genes for CLS have to be considered. The phenotypic similarities between CLS patients and *ld* mice remain striking, suggesting that CLS could well be caused by mutations in a gene whose product interacts with Gremlin or lies in the same pathway as Gremlin. It is possible that mutations in more than one gene are causing CLS and that once these genes are identified, they are found to be part of the same pathway and/or interact with each other in multi-component complexes.

The work on *Gremlin* being carried out in other laboratories, especially the studies of its function and the molecular pathway in which it acts, may help to suggest other

candidate genes to investigate in our panel of CLS patients. Moreover, a genome-wide screen in the probands from the four consanguineous families has to be considered in order to identify the CLS locus or loci.

Finally, clinicians should establish diagnostic criteria for CLS, taking into consideration the features described above, including renal malformation as well as facial dysmorphisms and rib and vertebral abnormalities.

Chapter 4 JATD results

4.1 Exclusion of candidate regions

4.1.1 *EVC/EVC2* region

In order to exclude the *EVC/EVC2* region as a possible candidate for JATD, a haplotype analysis was initially performed in patients IV.1 and IV.3 from Family 1, VI.3 and VI.5 from Family 2, II.1 from Family 3 and II.6 from Family 4. In a subsequent study, four further JATD families not consistent with linkage to any of the putative JATD loci subsequently identified were typed across the *EVC/EVC2* region. These were consanguineous Families 11 and 12, as well as non-consanguineous Families 13 and 15. It was thought that the parents in Family 13 were possibly distantly related. If this was the case therefore, and there was linkage to this locus, a homozygous region was expected to be found in patient III.1. In Family 15 there were three affected fetuses, therefore if there was linkage to this locus, a shared pair of haplotypes in all three patients was expected to be found in the region.

EVC and *EVC2* lie on the short arm of chromosome 4 (4p16) in a head-to-head configuration about 2 Kb apart (Fig 4.1) (Galdzicka *et al.*, 2002; Ruiz-Perez *et al.*, 2003). A total of 10 polymorphic markers were chosen to cover about 17 cM in the region of interest, but not all of them were run in each family. Dye-labelled primers for two markers (D4S412 and D4S2366) were taken from the linkage sets available in the laboratory (see section 2.3.2.1), while primers to a further four markers were chosen from the NCBI UniSTS database (see section 2.3.11.1) and designed as explained in section 2.3.2.2. Markers were chosen based on map position and degree of heterozygosity. Primers for marker D4S3360 were provided by collaborator Dr Colin Johnson. Primer sequences for markers 500H20P5, 500H20P3 (telomeric to *EVC2*) and *EVC*-exon 18 (CCCT repeat before *EVC* exon 18), were kindly provided by Prof. Judith Goodship to screen fetus II.2 in Family 12 (Ruiz-Perez *et al.*, 2003).

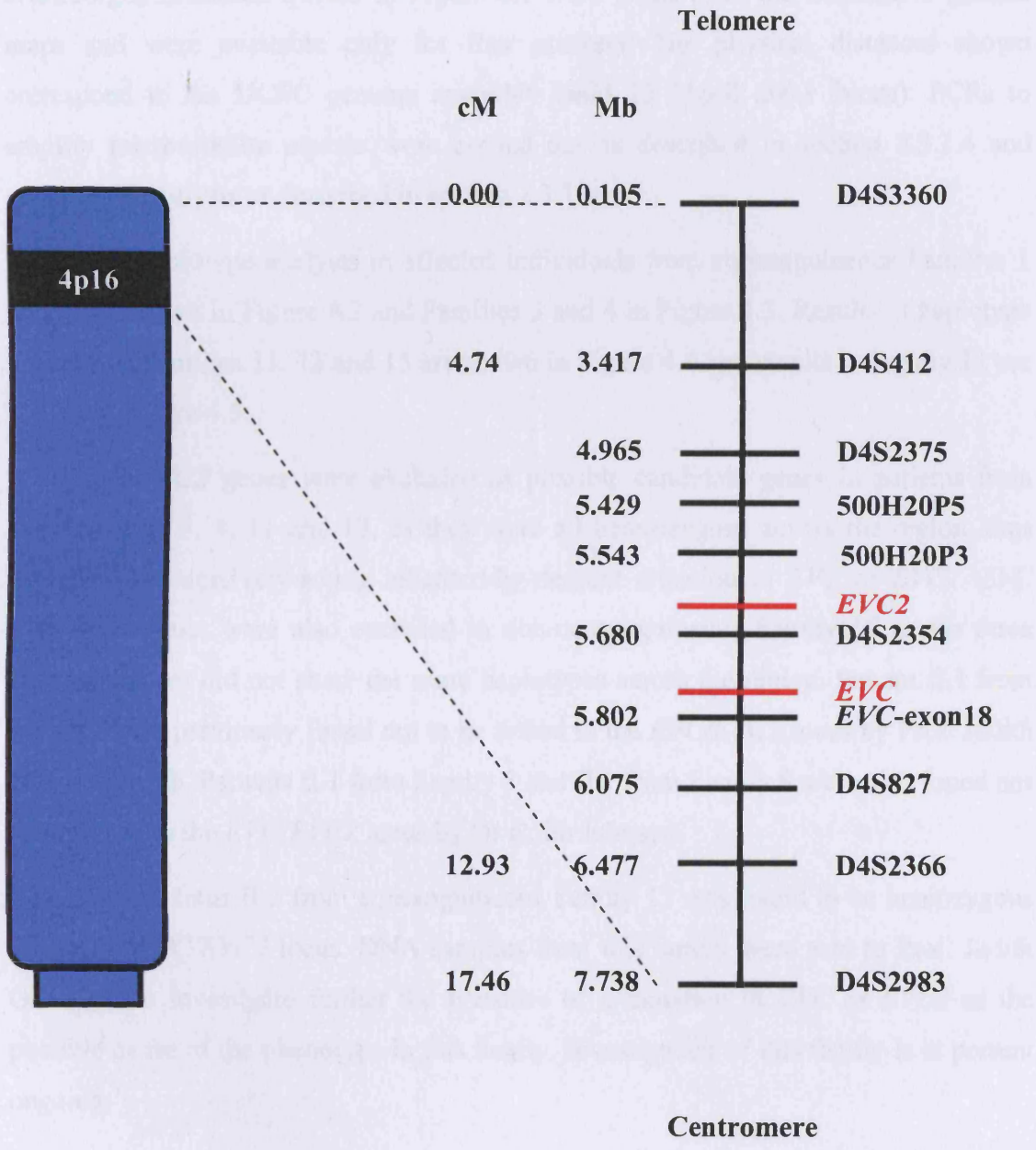


Figure 4.1 Schematic representation of the region on chromosome 4p16 where *EVC* and *EVC2* lie. Physical positions of polymorphic markers used for haplotype analysis are given for all markers.

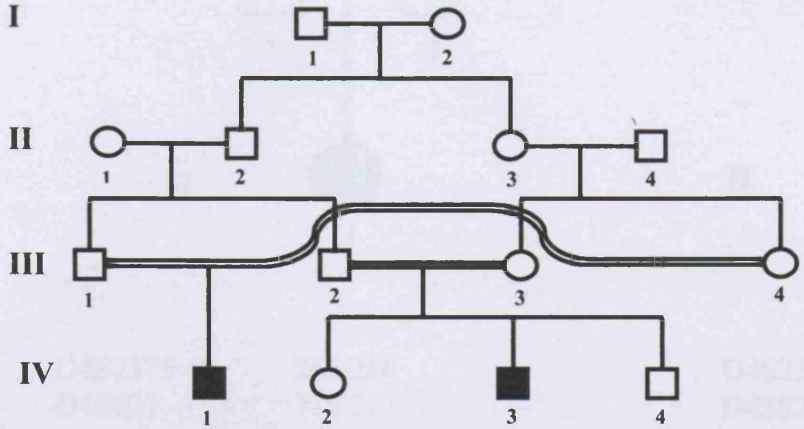
Details of the markers are shown in Table 2.2 and their positions in Figure 4.1. The centimorgan distances quoted in Figure 4.1 were taken from the Marshfield genetic maps and were available only for four markers. The physical distances shown correspond to the UCSC genome assembly build 33 (April 2003 freeze). PCRs to amplify microsatellite repeats were carried out as described in section 2.3.2.4 and genotyping analysis as described in section 2.3.7.1.

Results of haplotype analysis in affected individuals from consanguineous Families 1 and 2 are shown in Figure 4.2 and Families 3 and 4 in Figure 4.3. Results of haplotype analysis in Families 11, 13 and 15 are shown in Figure 4.4 and results in Family 12 are shown in Figure 4.5.

EVC and *EVC2* genes were excluded as possible candidate genes in patients from families 1, 2, 3, 4, 11 and 13, as they were all heterozygous across the region, thus excluding a recessively-acting inherited-by-descent mutation in *EVC* or *EVC2*. *EVC* and *EVC2* genes were also excluded in non-consanguineous Family 15 as the three affected fetuses did not share the same haplotypes across the region. Patient II.1 from Family 5 was previously found not to be linked to the *EVC/EVC2* locus by Prof. Judith Goodship's lab. Patients II.1 from Family 8 and II.1 from Family 9 were also found not to be linked to the *EVC/EVC2* locus by Dr Colin Johnson.

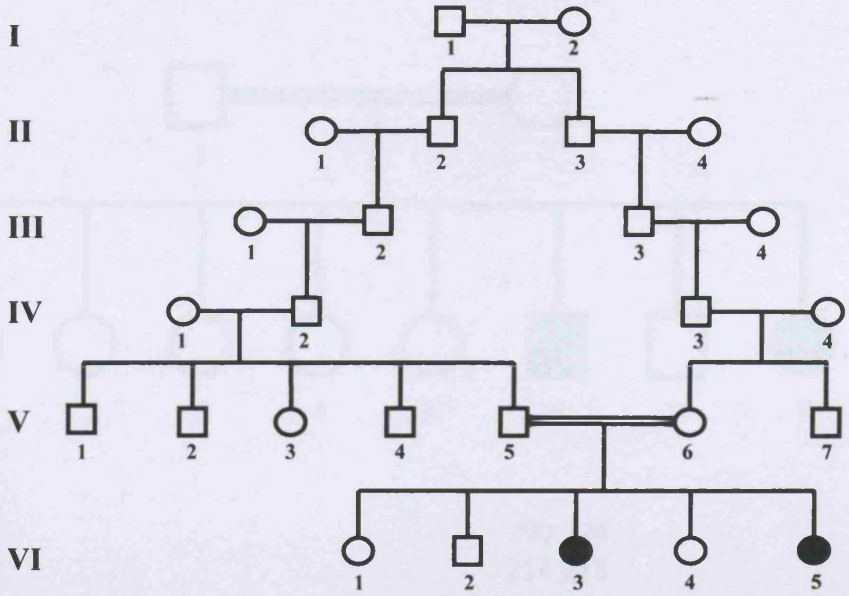
Nevertheless, fetus II.2 from consanguineous Family 12 was found to be homozygous across the *EVC/EVC2* locus. DNA samples from this family were sent to Prof. Judith Goodship to investigate further the presence of a mutation in *EVC* or *EVC2* as the possible cause of the phenotype in this family. Investigation of this family is at present ongoing.

Family 1



D4S3360	172 182	178 184
D4S412	162 164	156 166
D4S2375	286 286	266 286
D4S827	200 216	212 214
D4S2366	123 127	119 119
D4S2983	194 198	194 200

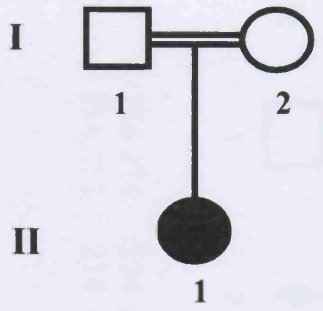
Family 2



D4S3360	172 174	172 178
D4S412	162 162	162 162
D4S2375	274 286	274 290
D4S827	214 218	204 214
D4S2366	123 135	123 123
D4S2983	190 194	190 194

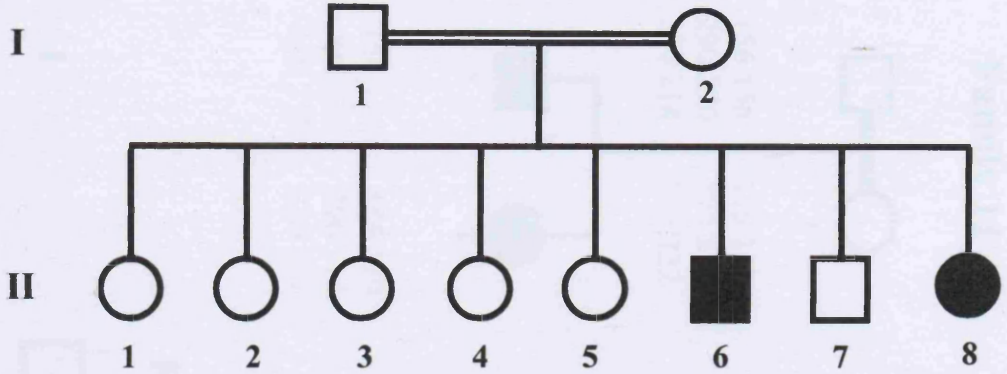
Figure 4.2 Results of haplotype analysis in affected individuals from Families 1 and 2 using 6 polymorphic markers on chromosome 4p16. Homozygous markers are highlighted in colour. None of the four affected individuals are homozygous across the *EVC/EVC2* locus. Only D4S412, about 2.1 Mb telomeric to the *EVC2* gene, is homozygous in both affected sibs from Family 2.

Family 3



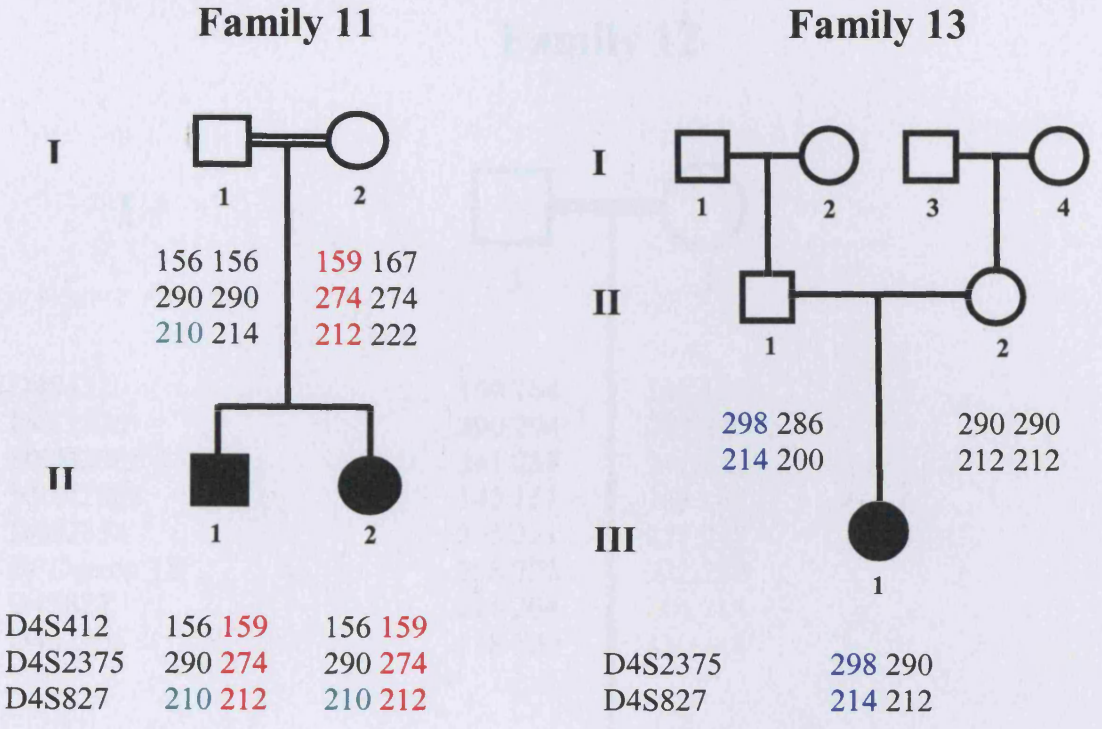
D4S2375	266 286
D4S827	214 214

Family 4



D4S2375	290 294
D4S827	214 218

Figure 4.3 Results of haplotype analysis in affected individuals from Families 3 and 4 using 2 polymorphic markers on chromosome 4p16. D4S2375 is located approximately 0.6 Mb telomeric to *EVC2*, while D4S827 is located 0.2 Mb centromeric to *EVC*. Homozygous result for D4S827 in patient II.1 (Family 3) is highlighted in red. Further investigation of this region in Family 3 was carried out by Dr Colin Johnson and linkage to the *EVC/EVC2* locus was excluded in this family.



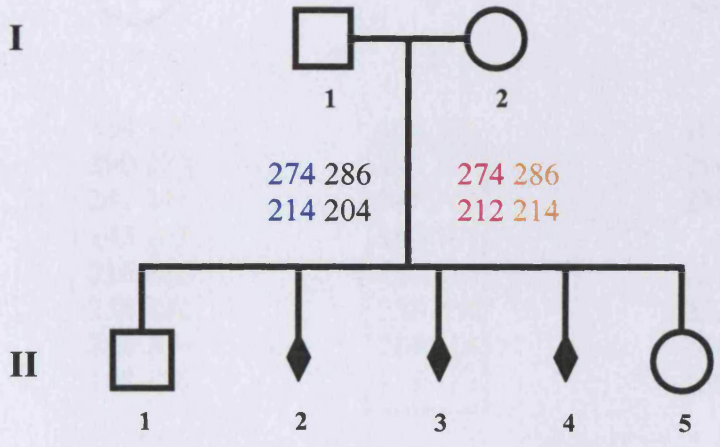
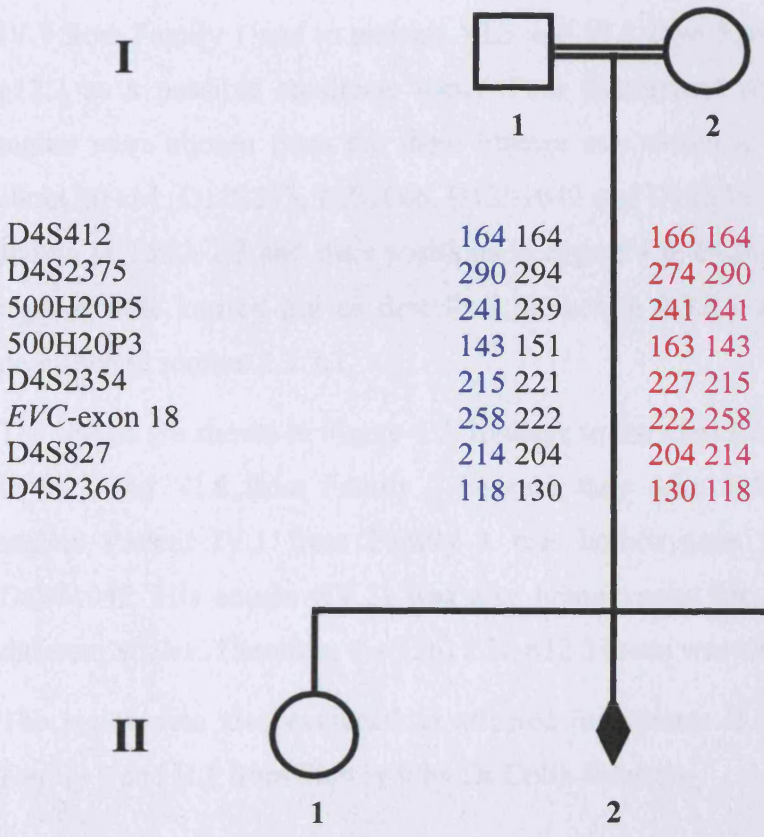


Figure 4.4 Results of haplotype analysis on chromosome 4p16 in Families 11, 13 and 15 showing no evidence of linkage to the *EVC/EVC2* locus. In Family 15, note the different paternal haplotype in fetus II.2 compared to fetuses II.3 and II.4 excluding linkage to the locus in this family.

Family 12



	I.1	I.2			
D4S412	164	164	166	164	
D4S2375	290	294	274	290	
500H20P5	241	239	241	241	
500H20P3	143	151	163	143	
D4S2354	215	221	227	215	
EVC-exon 18	258	222	222	258	
D4S827	214	204	204	214	
D4S2366	118	130	130	118	

	II.1		II.2		II.3	
D4S412	164	166	164	166	164	166
D4S2375	290	274	290	290	294	274
500H20P5	241	241	241	241	239	241
500H20P3	143	163	143	143	-	-
D4S2354	215	227	215	215	221	227
EVC-exon 18	258	222	258	258	222	222
D4S827	214	204	214	214	204	204
D4S2366	118	130	118	118	130	130

Figure 4.5 Haplotype analysis on chromosome 4p16 in Family 12 using 8 polymorphic markers in the region. Fetus II.2 demonstrates homozygosity from D4S2375 to D4S2366 (as shown by boxed haplotype) while the parents (I.1 and I.2) and two unaffected siblings (II.1 and II.3) are heterozygous across the region. Marker 500H20P3 could not be typed accurately in individual II.3 therefore the result is marked with a dash.

4.1.2 12p11-p12 region

A haplotype analysis was also performed in individuals III.1, III.2, III.3, III.4, IV.1 and IV.3 from Family 1 and in patients VI.3 and VI.5 from Family 2 to exclude 12p11.21-p12.2 as a possible candidate locus. Four fluorescent polymorphic markers in the region were chosen from the three linkage sets available in the laboratory to cover about 20 cM: D12S373, 12S1066, D12S1042 and D12S345. Details of the markers are shown in Table 2.3 and their positions in Figure 4.6. PCRs to amplified microsatellite repeats were carried out as described in section 2.3.2.4 and genotyping analysis as described in section 2.3.7.1.

The results are shown in Figure 4.7. Linkage to the 12p11.21-p12.2 locus was excluded in VI.3 and VI.5 from Family 2 because they were both heterozygous across this region. Patient IV.1 from Family 1 was homozygous for markers D12S373 and D12S1042. His cousin (IV.3) was also homozygous for markers D12S1042 but for different alleles. Therefore the 12p11.21-p12.2 locus was excluded in Family 1 as well.

The region was also excluded in affected individuals II.1 from Family 3, II.1 from Family 8 and II.1 from Family 9 by Dr Colin Johnson.

4.1.3 *Shorty* syntenic regions

The three human chromosomal regions syntenic to the *srt* locus were identified using the Human-Mouse Homology Map at http://www.ncbi.nlm.nih.gov/Omim/Homology (Herron et al., 2002). The 7 cM critical interval on mouse chromosome 17 between markers D17Mit134 and D17Mit51 corresponds to regions on human chromosomes 6p21, 6q25-q27 and 16p13.3. In order to exclude these loci as possible candidate regions for JATD, a haplotype analysis was performed in patients IV.1 and IV.3 from Family 1, VI.3 and VI.5 from Family 2. Four, five and three fluorescent polymorphic markers were chosen to cover the regions on 6p21, 6q25-q27 and 16p13.3 respectively. Six of them were available from the linkage sets in the laboratory (see section 2.3.2.1),

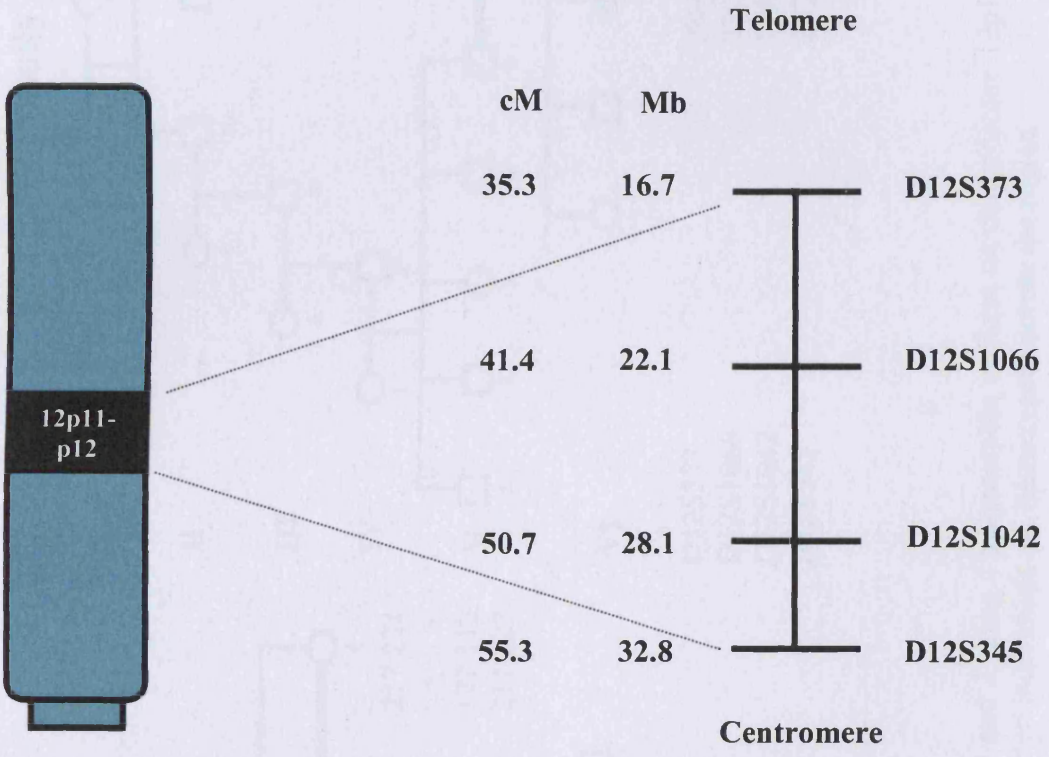
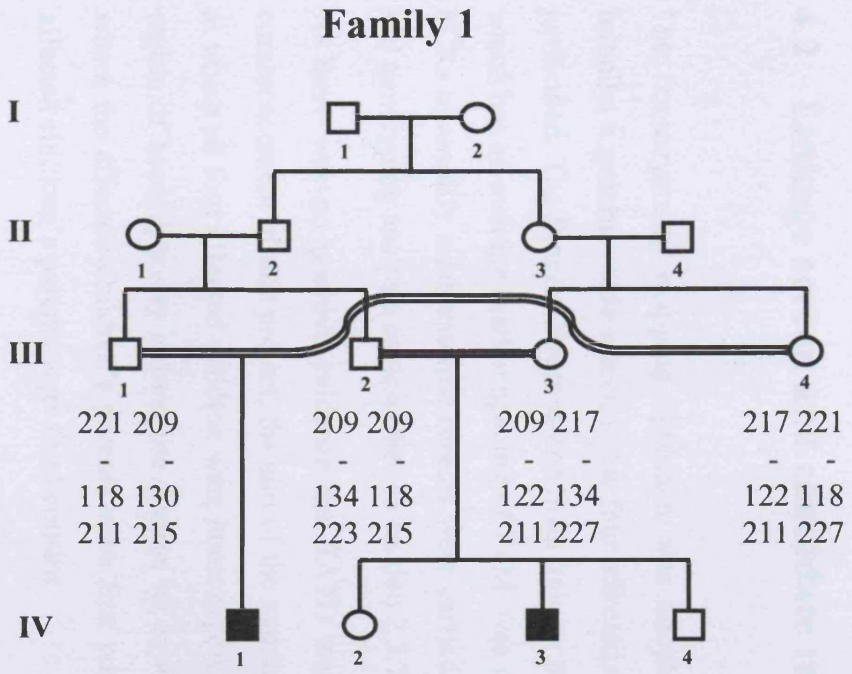
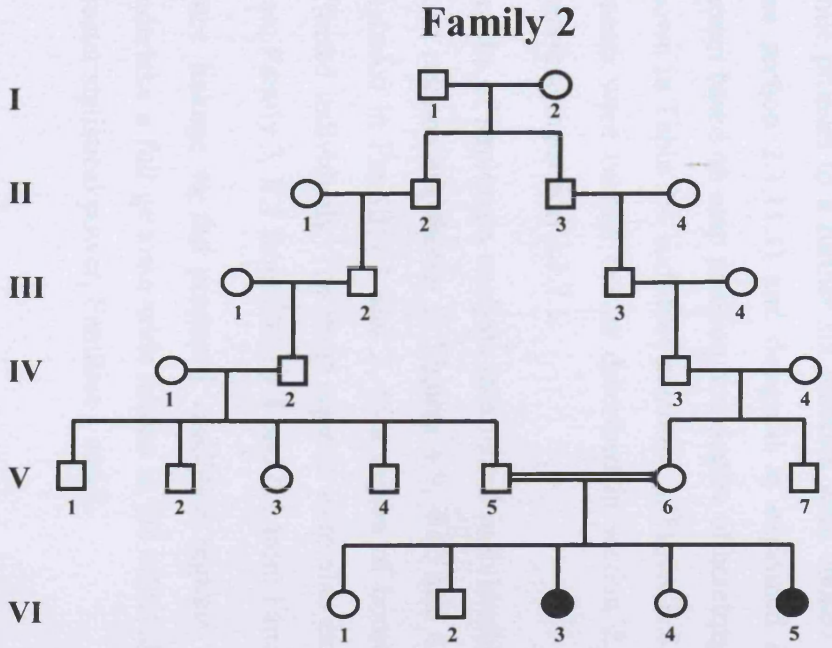


Figure 4.6 Schematic representation of the region on chromosome 12p11-p12. Genetic and physical positions of polymorphic markers used for haplotype analysis are given. The genetic distances quoted were taken from the high resolution deCode genetic map. The physical distances shown correspond to the NCBI genome assembly build 30.



D12S373	221 221	209 217	
D12S1066	210 214	214 222	
D12S1042	118 118	134 134	
D12S345	211 227	223 227	



D12S373		209 217	209 213
D12S1066		210 214	210 214
D12S1042		118 122	118 122
D12S345		213 215	211 215

Figure 4.7 Results of haplotype analysis in Families 1 and 2 using 4 polymorphic markers on chromosome 12p11-p12. Homozygous results are highlighted in colour. None of the four affected individuals are homozygous across the region.

while primers to a further six markers were chosen from the NCBI UniSTS database (see section 2.3.11.1) and designed as explained in section 2.3.2.2. Markers were chosen based on map position and degree of heterozygosity. Details of the markers are shown in Table 2.4 and their positions in Figure 4.8. PCRs to amplified microsatellite repeats were carried out as described in section 2.3.2.4 and genotyping analysis as described in section 2.3.7.1.

Results of haplotype analysis in affected individuals from Families 1 and 2 across all three regions are shown in Figures 4.9, 4.10 and 4.11. Linkage to all three loci was excluded in Families 1 and 2, as a region of homozygosity was not detected in the affected individuals. The three regions were also excluded in affected individuals II.1 from Family 3, II.1 from Family 8 and II.1 from Family 9 by Dr Colin Johnson.

Once linkage to the proposed candidate regions was excluded, it was decided to undertake a full genome-wide screen in the affected children of the two families with greater statistical power, Families 1 and 2.

4.2 Linkage screen and candidate regions

The homozygosity mapping approach was adopted to identify candidate regions. Initially, a genome-wide screen in the four affected children from Families 1 and 2 was performed. The Single Chromosome Scan linkage marker set from Research Genetics, which has an average marker spacing of 7 cM, was used as described in section 2.3.2.1. PCRs to amplify microsatellite repeats were carried out as described in section 2.3.2.4 and genotyping analysis as described in section 2.3.7.1.

As there was no positive evidence that JATD was genetically heterogeneous at the commencement of this project, the aim of the genome-wide screen was to find markers at which all four affected children were homozygously IBD. It was also thought that the region of homozygosity around the disease locus was likely to be larger in Family 1, where the affected children's parents were first cousins, than in Family 2, where the affected children's parents were third cousins.

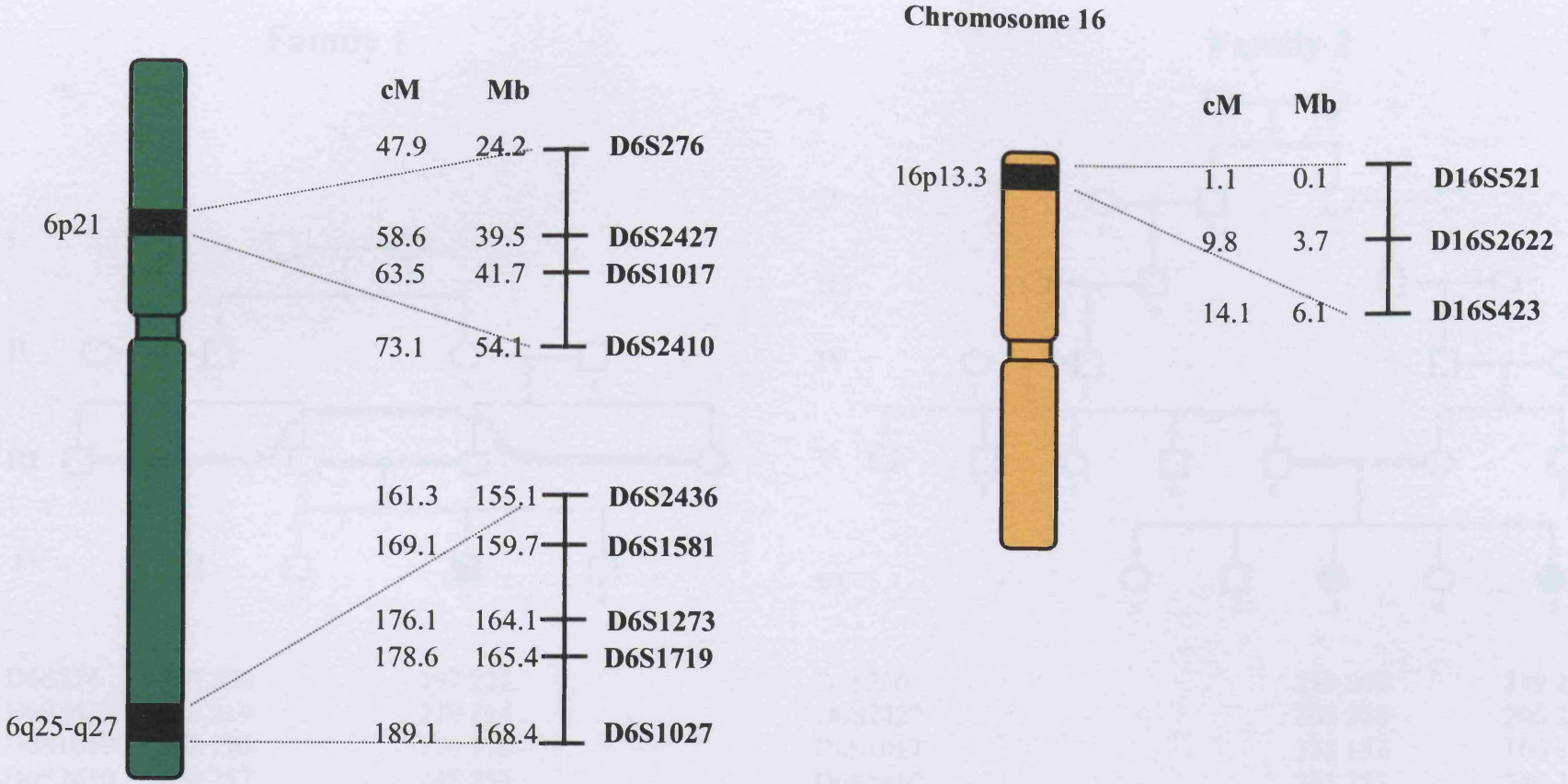
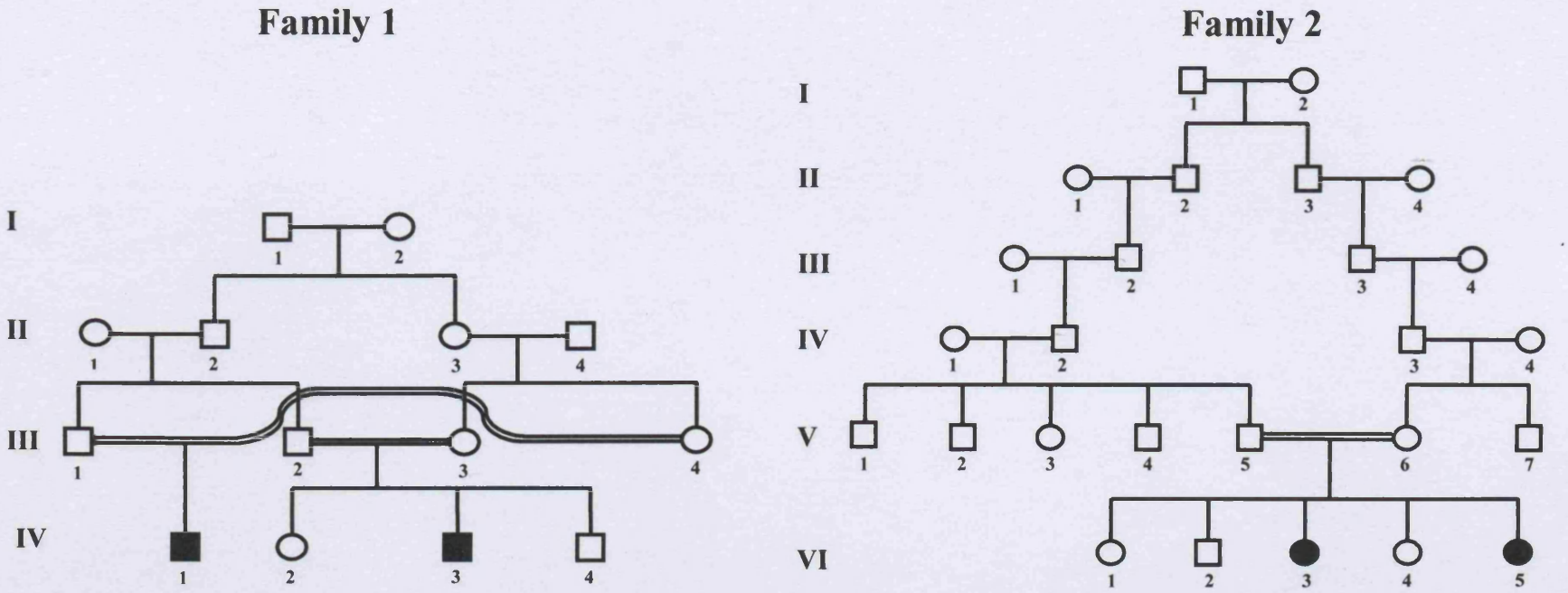


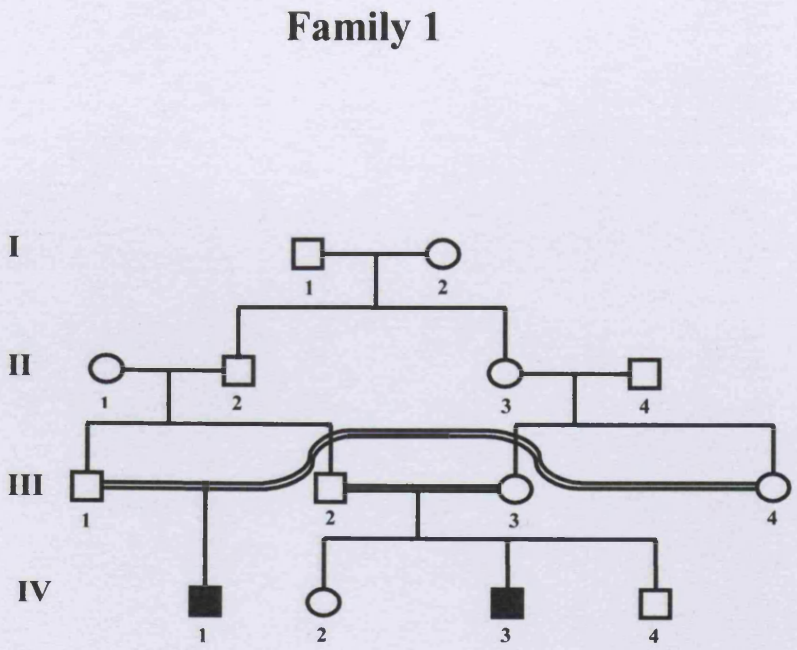
Figure 4.8 Schematic representation of the regions on chromosome 6p21, 6q25-q27 and 16p13.3 syntenic to the *srt* locus. Physical and genetic positions of polymorphic markers used for haplotype analysis are given for all markers. The genetic distances quoted were taken from the high resolution deCode genetic map. The physical distances listed correspond to the NCBI genome assembly build 30.



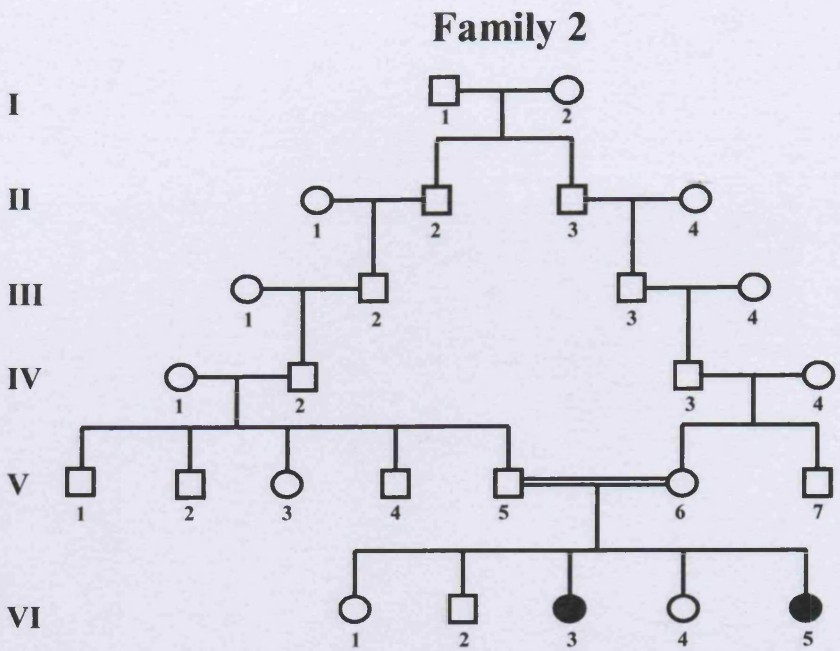
D6S276	207 209	197 221
D6S2427	210 210	210 214
D6S1017	172 176	176 176
D6S2410	249 257	245 253

D6S276	219 219	219 219
D6S2427	206 210	206 206
D6S1017	152 152	160 168
D6S2410	253 253	249 253

Figure 4.9 Results of haplotype analysis in affected individuals from Families 1 and 2 using 4 polymorphic markers on chromosome 6p21. Homozygous markers are highlighted in colour. None of the four affected individuals are homozygous across the locus. Only D6S276 is homozygous for the same alleles in both affected sibs from Family 2.



D6S2436	116 140	140 144
D6S1581	240 250	240 248
D6S1273	196 196	196 204
D6S1719	163 183	175 179
D6S1027	101 101	116 125



D6S2436	128 148	128 148
D6S1581	236 242	236 242
D6S1273	192 196	192 192
D6S1719	175 185	185 185
D6S1027	101 125	101 122

Figure 4.10 Results of haplotype analysis in affected individuals from Families 1 and 2 using 5 polymorphic markers on chromosome 6q25-q27. Homozygous markers are highlighted in colour. Note that patient VI.5 in Family 2 is homozygous for two consecutive markers (D6S1273 and D6S1719), but her affected sib VI.3 is heterozygous across the region. Linkage to this locus is excluded in both families.

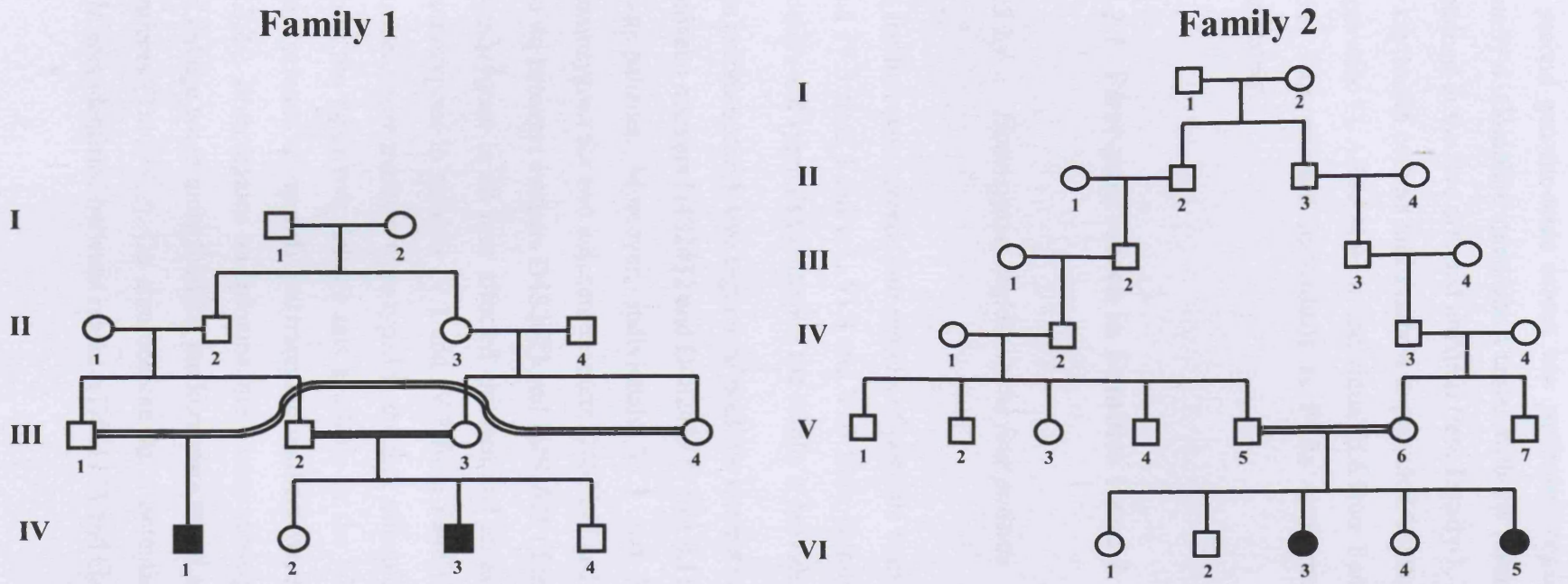


Figure 4.11 Results of haplotype analysis in affected individuals from Families 1 and 2 using 3 polymorphic markers on chromosome 16p13.3. Homozygous markers are highlighted in colour. Note that patient IV.1 in Family 1 is homozygous for two consecutive markers (D16S2622 and D16S423), but his affected cousin IV.3 is heterozygous. Linkage to this locus is excluded in both families.

A second genome-wide screen was performed separately in affected child II.6 from Family 4. Complete spreadsheet tables with the results of the full genome-wide screen obtained in the five affected children from Family 1, 2 and 4 are shown in Appendix 1. A key to the colours and symbols is provided to interpret the results (see first page of Appendix 1). Allele sizes in individual II.6 from Family 4 might be different from the other four affected individuals as PCRs were run in different gels and analysed separately.

4.2.1 First-pass screen in Families 1 and 2

4.2.1.1 Homozygous regions in the four patients

In the first-pass screen, four regions of possible homozygosity in all four patients (IV.1 and IV.3 from Family 1, VI.3 and VI.5 from Family 2) were identified. A key to the colours and symbols to interpret the results is provided in Figure 4.12.

On chromosome 4, two regions of possible interest were identified. The first was on 4q between markers D4S2432 and D4S2433 (Table 4.1). D4S3249 was homozygous in all four patients. Moreover, individuals VI.3 and VI.5 from Family 2 were also homozygous for two adjacent markers D4S3243 and D4S2361. The second region was on 4q between markers D4S2623 and D4S1625 (Table 4.1). D4S2394 was found to be homozygous in all four affected children and an adjacent marker D4S1644 was also homozygous in patients IV.1 and IV.3 from Family 1. To investigate the two regions further, more markers were typed in the four affected children. When possible, markers from the other two linkage sets available in the laboratory were used. Alternatively, new primers to amplify polymorphic markers were designed as explained in section 2.3.2.2. Both regions on chromosome 4 were subsequently excluded, with no evidence of linkage when more flanking markers were tested and found to be heterozygous in the patients (Table 4.2). On chromosome 8q, a potentially interesting region of about 30 cM was identified between markers D8S1113 and GAAT1A4 (Table 4.3).

Markers

Black (e.g. D8S1119) = 7 cM Single Chromosome Scan linkage marker set (Research Genetics, Inc.) – tri- and tetra-nucleotide repeats

Bold red (e.g. **D2S206**) = 10 cM ABI PRISM™ Linkage Mapping Set (Perkin Elmer Applied Biosystems Division) – di-nucleotide repeats

Bold purple (e.g. **D15S822**)= new markers chosen to fill in gaps – di-nucleotide repeats

Heterozygosity and distances

Het column gives heterozygosity of the marker available on the NCBI database uniSTS, Marshfield genetic map

Mb column gives distances in Mb from the top (tip of short arm) of the chromosome according to the UCSC Human Genome Assembly (April 2003 freeze) NCBI build 33

cM (Marshfield) column gives distances in cM from the top of the chromosome according to the Marshfield genetic map

cM (deCode) column gives distances in cM from the top of the chromosome according to the deCode genetic map

light blue fill indicates low heterozygosity

/ indicates heterozygosity or position unknown

Results

light green fill indicates individual is homozygous for that marker

dark green fill indicates all four patients are homozygous for that marker

green fill indicates two individuals from the same family are homozygous for 2 different alleles

bright green fill indicates candidate regions

light orange fill indicates result not clear

/ indicates no result could be obtained for that marker

Empty box indicates marker not tested

Figure 4.12 Key to genome-wide screen results tables.

Markers Chromosome 4	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5
D4S2366	0.77	6.47	12.93	12.46	123-127	119-119	123-135	123-123
D4S2639	0.88	18.53	/	/	175-183	163-175	159-183	159-175
GATA158G03	0.78	21.55	35.03	37.29	299-303	299-307	303-311	299-311
D4S2397	0.78	26.94	42.74		127-141	141-141	137-141	137-141
D4S3244	0.86	28.73	/	/	/	281-285	285-285	285-285
D4S2632	0.81	35.53	/	54.21	142-158	146-154	134-150	134-150
D4S1627	0.81	44.02	60.16	64.7	184-188	188-200	180-184	180-192
D4S2379	0.79	56.04	/	/	/	/	/	/
D4S1645	0.78	61.92	72.52	/	241-245	241-241	241-241	241-245
D4S2432	0.76	66.5	75.2	78.49	264-268	244-264	244-264	260-264
D4S2367	0.74	68.22	78.43	/	129-129	129-129	137-137	141-141
D4S3249	0.71	76.76	/	/	367-367	367-367	363-363	363-363
D4S3243	0.66	81.32	/	88.45	160-164	160-172	160-160	160-160
D4S2361	0.7	85.39	93.48	92.3	152-155	155-158	158-158	158-158
D4S2433	0.69	95.72	100.06	/	246-246	246-250	246-250	246-250
D4S1647	0.77	99.82	104.94	103.77	/	/	/	/
D4S2623	0.75	111.28	114.04	113.7	204-224	220-224	204-220	204-220
D4S2394	0.81	130.69	/	128.08	209-209	209-209	209-209	209-209
D4S1644	0.72	142.21	143.31	137.88	196-196	196-196	196-200	200-200
D4S1625	0.74	143.96	145.98	140.08	197-201	197-197	197-205	197-205
D4S1629	0.76	158.8	157.99	152.23	/	/	/	/
D4S2368	0.77	169.3	167.55	162.47	319-319	319-319	315-319	311-323
D4S2431	0.82	175.4	176.19	168.27	246-254	246-258	236-254	240-244
D4S2374	0.85	179.64	/	/	220-228	220-228	228-232	224-228
D4S2417	0.67	180.75	181.93	176.31	263-263	263-263	263-267	259-267
D4S3335	0.54	185.84	195.06	/	/	/	/	/
D4S2390	0.76	190.33	208.07	204.18	/	107-107	/	/

Table 4.1 Genotyping results on chromosome 4 obtained after the first pass genome-wide screen in patient IV.1, IV.3 from Family 1 and VI.3, VI.5 from Family 2 using microsatellite markers from the Single Chromosome Scan linkage marker set from Research Genetics. Heterozygosity, physical and genetic distances (both from the Marshfield and deCode genetic maps) of each marker are given. Arrows indicate the 2 markers at which all four patients were homozygous.

Markers Chromosome 4	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1		Family 2	
					IV.1	IV.3	VI.3	VI.5
D4S3360	/	0.1	0	0	172-182	178-184	172-174	172-178
D4S412	0.76	3.41	4.74	4.42	162-164	156-166	162-162	162-162
D4S431	0.83	6.4	12.35	12.46	154-158	158-162		
D4S2366	0.77	6.47	12.93	12.46	123-127	119-119	123-135	123-123
D4S2983	0.87	7.73	17.49	/	194-198	194-200	190-194	190-194
D4S403	0.76	13.43	25.9	26.71	179-179	169-179		
D4S2639	0.88	18.53	/	/	175-183	163-175	159-183	159-175
GATA158G03	0.78	21.55	35.03	37.29	299-303	299-307	303-311	299-311
D4S2397	0.78	26.94	42.74		127-141	141-141	137-141	137-141
D4S391	0.85	27.3	43.59	47.33	213-226	213-221		
D4S3244	0.86	28.73	/	/	/	281-285	285-285	285-285
D4S2632	0.81	35.53	/	54.21	142-158	146-154	134-150	134-150
D4S405	0.85	40.2	56.95	60.23	284-292	278-292		
D4S1627	0.81	44.02	60.16	64.7	184-188	188-200	180-184	180-192
D4S428	0.77	55.48	64.24	69.01	186-192	186-186		
D4S2379	0.79	56.04	/	/	/	/	/	/
D4S1645	0.78	61.92	72.52	/	241-245	241-241	241-241	241-245
D4S2432	0.76	66.5	75.2	78.49	264-268	244-264	244-264	260-264
D4S2367	0.74	68.22	78.43	/	129-129	129-129	137-137	141-141
D4S392	0.83	70.77	78.97	80.38	101-103	101-101	99-105	99-103
D4S3249	0.71	76.76	/	/	367-367	367-367	363-363	363-363
D4S3042	0.84	77.29	83.29	85.42	196-214	196-218		
D4S2964	0.77	81.16	88.35	88.45	237-237	237-245	225-237	225-237
D4S3243	0.66	81.32	/	88.45	160-164	160-172	160-160	160-160
D4S2361	0.7	85.39	93.48	92.3	152-155	155-158	158-158	158-158
D4S2433	0.69	95.72	100.06	/	246-246	246-250	246-250	246-250
D4S414	0.89	92.83	100.75	100.16	227-233	233-233		
D4S1647	0.77	99.82	104.94	103.77	/	/	/	/

D4S1572	**0.84**	**104.16**	**107.95**	**107.52**	200-202	200-202	198-198	198-202
D4S2623	**0.75**	**111.28**	**114.04**	**113.7**	204-224	220-224	204-220	204-220
D4S402	**0.9**	**120.54**	**124.45**	**122.1**	129-135	131-135	107-127	107-109
D4S2938	**0.78**	**129.71**	**129.38**	/	183-183	179-179	177-177	187-187
D4S2394	**0.81**	**130.69**	/	**128.08**	209-209	209-209	209-209	209-209
D4S1575	**0.5**	**135.18**	**132.05**	**131.59**	290-290	290-290	288-290	290-290
D4S3039	**0.76**	**136.35**	**132.72**	/	166-172	168-172		
D4S1644	**0.72**	**142.21**	**143.31**	**137.88**	196-196	196-196	196-200	200-200
D4S424	**0.83**	**142.66**	**144.56**	**138.87**	194-198	198-204		
D4S1625	**0.74**	**143.96**	**145.98**	**140.08**	197-201	197-197	197-205	197-205
D4S1629	**0.76**	**158.8**	**157.99**	**152.23**	/	/	/	/
D4S413	**0.86**	**158.81**	**157.99**	/	282-282	282-294	285-318	293-316
D4S2368	**0.77**	**169.3**	**167.55**	**162.47**	319-319	319-319	315-319	311-323
D4S1597	**0.74**	**170.42**	**169.42**	**163.65**	274-290	/		
D4S2431	**0.82**	**175.4**	**176.19**	**168.27**	246-254	246-258	236-254	240-244
D4S2374	**0.85**	**179.64**	/	/	220-228	220-228	228-232	224-228
D4S2417	**0.67**	**180.75**	**181.93**	**176.31**	263-263	263-263	263-267	259-267
D4S1535	**0.75**	**185.82**	**195.06**	**189.38**	247-253	251-259	253-259	251-253
D4S3335	**0.54**	**185.84**	**195.06**	/	/	/	/	/
D4S426	**0.78**	**189.69**	**206.98**	**202.69**	160-160	156-156	158-158	150-170
D4S2390	**0.76**	**190.33**	**208.07**	**204.18**	/	107-107	/	/

Table 4.2 Genotyping results on chromosome 4 obtained in patient IV.1, IV.3 from Family 1 and VI.3, VI.5 from Family 2 after testing more microsatellite markers. Note that regions between markers D4S2432 and D4S2433 and between D4S2623 and D4S1625 (indicated by brackets) no longer represent blocks of homozygosity in the four patients.

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1		Family 2	
					IV.1	IV.3	VI.3	VI.5
D8S1469	0.67	8.96	16.19	/	/	213-213	193-209	205-209
D8S1130	0.8	11.7	22.41	18.18	/	134-138	134-142	142-146
D8S1106	0.73	12.61	26.43	23.58	137-141	141-141	137-141	/
D8S1989	0.61	24.46	50.05	/	/	/	211-215	193-211
D8S1477	0.86	31.92	60.34	/	162-162	162-166	138-162	138-138
D8S1104	0.68	40.38	64.6	59.17	132-136	136-140	128-132	128-140
D8S1110	0.77	52.9	67.27	65.47	280-280	260-280	272-276	276-276
D8S1113	0.81	59.43	77.89	70.74	232-236	232-236	220-236	216-236
D8S1136	0.74	65.79	82.26	/	/	253-253	245-257	249-257
D8S2323	0.84	69.79	87.54	/	302-302	302-302	298-302	302-302
GATA14E09	0.75	73.97	94.08	/	204-204	204-204	204-212	204-212
D8S1119	0.8	86.84	101.01	/	173-173	173-173	181-181	181-181
D8S1988	0.63	93.15	104.33	97.63	/	/	/	/
GAAT1A4	0.66	98.85	110.2	104.28	148-156	156-156	152-156	152-156
D8S1459	0.68	105.08	117.62	/	125-131	125-131	131-134	131-134
D8S1470	0.81	111.89	122.96	115.83	354-358	354-362	354-362	354-362
D8S1471	0.73	113.37	123.54	116.22	289-289	285-297	293-293	293-293
D8S1142	0.69	114.42	/	/	354-358	350-354	354-354	354-354
D8S592	0.67	118.12	125.27	118.41	165-165	161-173	161-173	161-173
D8S586	0.85	120.85	128.16	121.37	/	/	/	/
D8S1179	0.81	125.57	135.08	129.33	173-185	181-185	169-181	169-181
D8S568	0.82	126.4	/	/	254-254	254-270	258-266	258-266
D8S1128	0.76	128.26	139.53	135.57	/	/	/	238-242
D8S1462	0.75	135.06	/	/	156-160	156-156	160-160	160-160
D8S1100	0.65	136.77	154.02	148.81	185-188	188-188	182-182	182-182

Table 4.3 Genotyping results on chromosome 8 obtained after the first pass genome-wide screen in patient IV.1, IV.3 from Family 1 and VI.3, VI.5 from Family 2 using microsatellite markers from the Single Chromosome Scan linkage marker set from Research Genetics. Marker D8S1119 is homozygous in all four patients and a potential shared region of homozygosity in patients IV.1 and IV.3 from Family 1 between markers D8S1113 and GAAT1A4 (indicated by bracket) could extend up to 30 cM.

This region of homozygosity was subsequently excluded when more markers known to lie in the region were tested and were found to be heterozygous (Table 4.4). In addition marker D8S1119 was run on separate gel lanes for the four patients and was found to be heterozygous in individuals IV.1 and IV.3 from Family 1. Marker GATA14E09 was also found to be heterozygous in individual IV.3 from Family 1 when the sample was run on a separate gel lane (Table 4.4).

After the first-pass screen, all four patients were homozygous for marker D17S1298 located at 10.72 cM from the 17pter (See Appendix 1). When a very close marker D17S1876 was typed in all four affected children, individual IV.1 from Family 1, VI.3 and VI.5 from Family 2 were found to be heterozygous, thereby excluding the region.

4.2.1.2 Difficulties of the genome-wide screen

In general, the results from the first-pass screen were unsatisfactory. They gave inadequate coverage of the genome. There were several reasons for this. First of all the Research Genetics linkage set contains a certain number of gaps between markers; three gaps greater than 20 cM, fourteen gaps greater than 15 cM, and seventy greater than 10 cM were counted. Therefore, even if all markers had worked perfectly, complete coverage of the genome could not have been achieved. Furthermore, other gaps were generated where the PCR failed to amplify the markers. Out of 516 pairs of primers, 55 (10.6%) did not work at all the first time and 70 (13.5%) worked poorly with at least one reaction out of four that did not work. For markers that did not work the first time, the PCR conditions were modified. Lowering the annealing temperature to 53°C resolved the problem in some cases; loading more PCR product gave good results, especially for the faint, HEX-labelled products.

For markers with good PCR product, results were often hard to interpret. Sometimes the presence of extra peaks due to non-specific bands in the PCR or superimposed peaks from other colours made the results confusing. In order to avoid these problems, PCR was repeated and PCR products checked on a 2% agarose gel to ensure that only a single specific band was amplified. PCR conditions were changed (increasing the

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5
D8S504	0.72	1.16	0	0	139-147	139-139	131-137	137-141
D8S1469	0.67	8.96	16.19	/	/	213-213	193-209	205-209
D8S1130	0.8	11.7	22.41	18.18	/	134-138	134-142	142-146
D8S1106	0.73	12.61	26.43	23.58	137-141	141-141	137-141	/
D8S258	0.7	20.17	41.55	34.88	150-152	146-150	146-150	146-150
D8S1989	0.61	24.46	50.05	/	/	/	211-215	193-211
D8S1477	0.86	31.92	60.34	/	162-162	162-166	138-162	138-138
D8S1104	0.68	40.38	64.6	59.17	132-136	136-140	128-132	128-140
D8S1110	0.77	52.9	67.27	65.47	280-280	260-280	272-276	276-276
D8S1113	0.81	59.43	77.89	70.74	232-236	232-236	220-236	216-236
D8S260	0.82	61.54	79.36	73.6	204-204	198-208	204-204	204-204
D8S1136	0.74	65.79	82.26	/	/	253-253	245-257	249-257
D8S1797	0.83	66.97	82.84	77.01	145-145	133-143		
D8S2323	0.84	69.79	87.54	/	302-302	302-302	298-302	302-302
D8S279	0.88	72.71	91.46	84.3	238-244	238-244	230-246	230-246
GATA14E09	0.75	73.97	94.08	/	204-204	204-216	204-212	204-212
D8S1760	0.81	80.27	96.21	/	140-152	152-154	154-156	154-156
D8S1119	0.8	86.84	101.01	/	172-187	172-181	181-181	181-181
D8S1988	0.63	93.15	104.33	97.63	291-291	287-295	287-291	287-291
GAAT1A4	0.66	98.85	110.2	104.28	148-156	156-156	152-156	152-156
D8S1459	0.68	105.08	117.62	/	125-131	125-131	131-134	131-134
D8S1470	0.81	111.89	122.96	115.83	354-358	354-362	354-362	354-362
D8S1471	0.73	113.37	123.54	116.22	289-289	285-297	293-293	293-293
D8S1142	0.69	114.42	/	/	354-358	350-354	354-354	354-354
D8S592	0.67	118.12	125.27	118.41	165-165	161-173	161-173	161-173
D8S586	0.85	120.85	128.16	121.37	245-245	245-249	/	245-249
D8S514	0.76	123.41	130	124.62	216-218	218-222	222-222	214-222
D8S1179	0.81	125.57	135.08	129.33	173-185	181-185	169-181	169-181
D8S568	0.82	126.4	/	/	254-254	254-270	258-266	258-266

D8S1128	0.76	128.26	139.53	135.57	/	/	/	238-242
D8S284	0.83	131.18	143.82	139.79	278-296	270-296	296-296	296-296
D8S1462	0.75	135.06	/	/	156-160	156-156	160-160	160-160
D8S1100	0.65	136.77	154.02	148.81	185-188	188-188	182-182	182-182
D8S272	0.8	137.41	154.02	/			253-253	253-253

Table 4.4 Genotyping results on chromosome 8 after testing more polymorphic markers. The region between markers D8S1113 and GAAT1A4 (indicated by bracket) is no longer homozygous in all four patients. Note different results for markers GATA14E09 and D8S1119 after repeating the PCRs and running the samples separately in single gel lanes.

temperature) in the case of non-specific bands. Samples were then run in separate lanes on the ABI gel to avoid peaks coming through from other colours. Another problem that was encountered was that the range of allele sizes in a panel was sometimes too close, resulting in overlapping peaks within the same Genotyper window, again confusing the results. Finally, the range of allele sizes was often different from the one stated by the manufacturers. Genotyper calls peaks only in the interval of allele size that is given. If that interval is not accurate, one risks missing peaks and calling patients homozygous, when they are actually heterozygous.

4.2.2 Second-pass screen

In order to resolve some of the problems discussed above, a new strategy was adopted. In order to get an adequate coverage of the genome (with no gaps larger than 10 cM), markers from the other two linkage sets available in the laboratory were used and/or new primers were designed. The problem with using the two other linkage sets was that they were both quite old and thus some primers were missing from their boxes, while others were not working properly. In the latter case, it was necessary to vary the PCR conditions to amplify the products. In addition, it was difficult to define exactly the physical and genetic positions of the markers, especially for the Research Genetics markers. Some markers were not listed at all in the uniSTS database and some gave only the position in cM. For some of these, the position in Mb was obtained by doing a BLAST search (http://www.ncbi.nlm.nih.gov/BLAST/) using the sequence surrounding the marker against the human genome sequence.

Therefore, new microsatellite markers were ascertained from both Map View at NCBI (http://www.ncbi.nlm.nih.gov/mapview) and the UCSC Human Genome Browser at http://www.genome.ucsc.edu. Both databases were used to integrate genetic and physical map data. In the initial stages of the project, linkage map data from the Marshfield Centre for Medical Genetics was primarily used to locate markers, their primer sequences and heterozygosities. Later, both the deCode Icelandic map (available on the NCBI database, uniSTS) and the UCSC database, were preferred to choose

markers. A second strategy was to search for novel polymorphic markers by BLAST searching each BAC in a particular region with a nucleotide repeat sequence, such as $ca^{(n)}$ or $tg^{(n)}$. Although not all the repeats identified with this strategy were found to be polymorphic, a few were successfully used for linkage because of their high heterozygosity.

The aim of the second-pass screen was to fill in any gaps greater than 7-10 cM between the markers tested. After this second-pass screen, a number of regions in which all four patients were homozygous for one or more markers were identified (Appendix 1). To check whether the four children were really IBD at these markers, or just IBS, genotyping with flanking markers, and when necessary genotyping of their parents and siblings, was carried out.

A possibly interesting region, for example, was located on chromosome 17q25 between marker D17S968 and D17S928 (Table 4.5A). All four patients were homozygous for D17S801 and patients IV.1 and IV.3 from Family 1 were also homozygous for the three adjacent markers D17S802, D17S1847 and D17S784 (Table 4.5A). After more markers in the region were typed, the region was excluded in both families (Table 4.5B). Furthermore, haplotype analysis of the parents and the children in the region was performed and they were all found to be homozygous for D17S801 (Fig. 4.13 and 4.14).

4.2.3 Identification of candidate regions

After the second-pass screen, no region that was IBD in all four patients was identified. A possible explanation for this is locus heterogeneity, namely that the disease loci in Families 1 and 2 were different. A number of homozygous regions were found in each family individually and each of them was investigated further. Because the region of homozygosity around the disease locus was likely to be larger in Family 1 than in Family 2, initially it was decided to focus on defining and fine-mapping regions that were IBD in the two affected children from Family 1.

A)

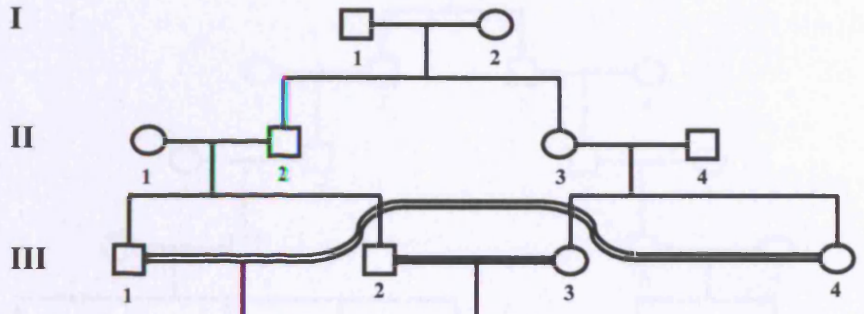
Markers Chromosome 17	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5
ATA43A10	0.79	67.01	89.32	99.34	146-150	137-150	154-162	154-162
D17S968	/	73.07	/	/	162-170	162-170	166-170	170-170
D17S801	0.85	74.97	103.53	/	269-269	269-269	269-269	269-269
D17S802	0.81	76.7	106.8	120.84	185-185	185-185	181-183	181-183
D17S1847	0.66	77.48	111.22	123.75	268-268	268-268	266-268	266-268
D17S784	0.77	78.36	116.86	129.62	231-231	231-231	235-235	235-235
D17S928	0.77	80.78	126.46	135.67	81-87	81-87	84-87	84-87

B)

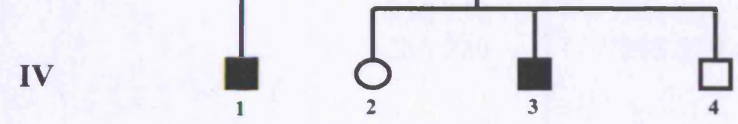
Markers Chromosome 17	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5
ATA43A10	0.79	67.01	89.32	99.34	146-150	137-150	154-162	154-162
D17S968	/	73.07	/	/	162-170	162-170	166-170	170-170
D17S1603	0.73	74.53	102.99	115.06	227-241	227-241	226-228	228-234
D17S785	0.83	74.89	103.53	115.34	191-205	191-205		
D17S801	0.85	74.97	103.53	/	269-269	269-269	269-269	269-269
D17S802	0.81	76.7	106.8	120.84	185-185	185-185	181-183	181-183
D17S1847	0.66	77.48	111.22	123.75	268-268	268-268	266-268	266-268
D17S836	0.63	77.76	112.92	125.04	201-205	201-205	201-205	201-205
D17S1806	0.72	77.9	114.41	/	247-259	247-259	253-253	253-253
D17S1822	0.7	78.34	116.86	129.24	250-252	250-252	244-250	244-250
D17S784	0.77	78.36	116.86	129.62	231-231	231-231	235-235	235-235
D17S928	0.77	80.78	126.46	135.67	81-87	81-87	84-87	84-87

Table 4.5 Genotyping results on chromosome 17q25 in patient IV.1, IV.3 from Family 1 and VI.3, VI.5 from Family 2 before (A) and after (B) the second-pass screen showing how a potentially interesting region was excluded by running more polymorphic markers. D17S785, very close to D17S801, is heterozygous in IV.1 and IV.3 from Family 1. D17S836, very close to D17S1847, is heterozygous in all four patients.

Family 1



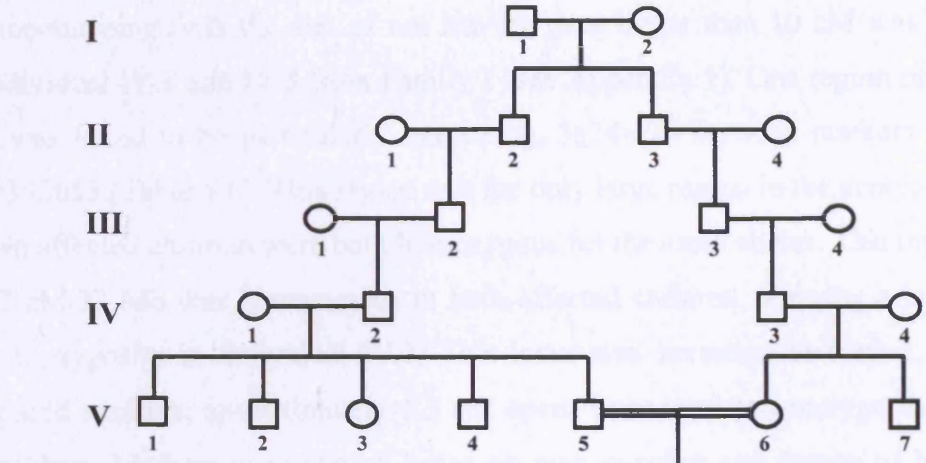
	III.1	III.2	III.3	III.4
D17S801	269 269	269 269	269 269	269 269
D17S802	185 187	185 187	185 187	185 187
D17S1847	268 270	268 268	268 268	268 268
D17S836	201 205	201 205	205 205	205 205
D17S1806	247 253	247 253	259 253	259 247
D17S1822	250 250	250 250	252 252	252 250
D17S784	231 233	231 233	231 237	231 233



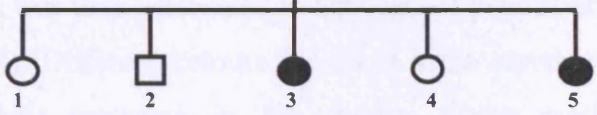
	IV.1	IV.3
D17S801	269 269	269 269
D17S802	185 185	185 185
D17S1847	268 268	268 268
D17S836	201 205	201 205
D17S1806	247 259	247 259
D17S1822	250 252	250 252
D17S784	231 231	231 231

Figure 4.13 Results of haplotype analysis in Family 1 using seven polymorphic markers on chromosome 17q25. Homozygous results are highlighted in colour. Note that marker D17S801 is homozygous in the two affected children, but also in their parents. D17S1847 is semi-informative. D17S784 is homozygous in patients IV.1 and IV.3, but a very close marker, D17S1822, is not.

Family 2



D17S801	- -		269 269
D17S802	- -		- -
D17S1847	- -		- -
D17S836	201 205		205 205
D17S1806	253 255		253 245
D17S1822	250 252		244 244
D17S784	235 229		235 229



D17S801	269 269	269 269	269 269
D17S802	- -	181 183	181 183
D17S1847	- -	266 268	266 268
D17S836	205 205	201 205	201 205
D17S1806	255 253	253 253	253 253
D17S1822	252 244	250 244	250 244
D17S784	229 235	235 235	235 235

Figure 4.14 Results of haplotype analysis in Family 2 using seven polymorphic markers on chromosome 17q25. Homozygous results are highlighted in colour. Note that marker D17S801 is homozygous in the two affected children, but also in their parents and their unaffected brother. D17S784 is homozygous in patients VI.3 and VI.5, but a very close marker, D17S1822, is not. Markers which were not typed for a particular individual are marked with a dash.

4.2.3.1 Family 1

Fine-mapping with the aim of not leaving gaps larger than 10 cM was carried out in individual IV.1 and IV.3 from Family 1 (see Appendix 1). One region on chromosome 3 was found to be particularly interesting, 3q24-q26 between markers D3S1764 and D3S3053 (Table 4.6). This region was the only large region in the genome in which the two affected children were both homozygous for the same alleles. This interval of about 17 cM/23 Mb was homozygous in both affected children, showing a larger region of homozygosity in individual IV.1. This locus was investigated further. More closely spaced markers, approximately 1-3 cM apart, were used to genotype the two affected children. Markers were chosen based on map position and degree of heterozygosity. Primers to amplify polymorphic markers were designed as described in section 2.3.2.2. Details of markers are shown in Table 2.5 and their positions in Table 4.7. Markers R1, R2 and R3 were found by searching for repeats in three BAC clones between markers D3S3723 and D3S1574 in order to better refine the minimal critical interval.

The two affected children continued to show homozygosity over the whole region (Table 4.7). Haplotype analysis was then performed in both sets of parents as well as in unaffected sib IV.2 (Table 4.8). Different colours for allele sizes represent different haplotypes to show how they segregate in the family. Some markers were uninformative or semi-informative. In this case, the allele sizes were left in black for the children. Both sets of parents and unaffected sib IV.2 were heterozygous for most of the markers supporting the hypothesis of linkage to this locus. The minimal critical interval was defined between markers D3S1555 and D3S1308 at the centromeric end due to a cross-over in individuals III.2 and between markers R3 and D3S1574 at the telomeric end due to a paternal cross-over in patient IV.3 (Table 4.8).

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3
D3S2406	0.88	73.06	102.64	/	315-327	327-339
D3S4529	0.72	85.65	112.42	/	158-158	154-158
D3S1271	0.74	102.01	117.76	112.28	93-93	85-93
D3S2459	0.84	103.45	119.09	113.38	189-189	185-201
D3S3045	0.82	108.27	124.16	117.29	179-179	175-179
D3S4018	0.78	112.93	127.89	120.11	292-292	292-300
D3S2460	0.76	118.68	134.64	/	149-149	161-161
ATA34G06	0.71	/	138	126.98	230-230	230-242
D3S1267	0.88	124.32	139.12	/	93-93	93-95
D3S4527	0.7	128.17	/	/	239-239	227-235
D3S1292	0.84	132.91	146.6	138.82	130-130	116-128
D3S1764	0.79	140.46	152.62	145.53	241-241	241-245
D3S1569	0.8	144.65	158.38	150.58	152-152	152-152
D3S1744	0.77	148.37	161.04	153.36	150-150	150-150
D3S1279	0.85	152.3	169.6	160.19	266-266	266-266
D3S4531	0.69	152.53	169.6	160.46	263-263	263-263
D3S1763	0.8	168.64	176.54	168.52	277-277	277-277
D3S1614	0.83	169.61	177.75	169.98	110-110	110-110
D3S3053	0.73	173.15	181.87	/	232-232	236-232
D3S1565	0.63	174.88	186.04	177.94	182-182	182-182
D3S2427	0.85	177.18	188.29	/	230-230	236-230
D3S1262	0.8	187.62	201.14	194.33	114-114	114-114
D3S2436	0.67	188.23	203.28	/	171-171	171-171
D3S1580	0.84	189.94	207.73	202	224-218	218-218
D3S2398	0.77	190.82	209.41	203.31	274-294	294-294
D3S2418	0.71	193.72	215.84	/	106-109	106-109

Table 4.6 Chromosome 3q24-q26 results obtained in patients IV.1 and IV.3 from Family 1 after the second-pass screen showing a region of shared homozygosity in the two affected individuals between markers D3S1764 and D3S3053.

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3
D3S2406	0.88	73.06	102.64	/	315-327	327-339
D3S4529	0.72	85.65	112.42	/	158-158	154-158
D3S1271	0.74	102.01	117.76	112.28	93-93	85-93
D3S1764	0.79	140.46	152.62	145.53	241-241	241-245
D3S1309	0.74	142	153.74	146.63	167-167	171-165
D3S3694	0.84	143.47	153.74	148.82	302-302	294-300
D3S1569	0.8	144.65	158.38	150.58	152-152	152-152
D3S1593	0.78	146.6	161.04	152.31	226-226	212-226
D3S1744	0.77	148.37	161.04	153.36	150-150	150-150
D3S1306	0.72	149.08	164.25	154.18	158-158	164-158
D3S1555	0.8	150.08	165.32	156.59	211-211	203-211
D3S1308	0.69	150.31	165.85	157.07	190-190	190-190
D3S1299	0.62	151.46	166.93	158.56	196-196	196-196
D3S1279	0.85	152.3	169.6	160.19	266-266	266-266
D3S4531	0.69	152.53	169.6	160.46	263-263	263-263
D3S1746	0.85	153.01	169.6	/	271-271	271-271
D3S1280	0.75	153.38	170.14	161.6	210-210	210-210
D3S3531	0.73	155.49	170.14	162.79	218-218	218-218
D3S1607	0.8	158.24	172.27	164.61	235-235	235-235
D3S3575	0.79	160.63	173.34	165.61	188-188	188-188
D3S3579	0.79	161.86	173.87	166.45	186-186	186-186
D3S1268	0.86	165.29	174.94	/	204-204	204-204
D3S3668	0.83	165.72	175.47	167.47	228-228	228-228
D3S1763	0.8	168.64	176.54	168.52	277-277	277-277
D3S1614	0.83	169.61	177.75	169.98	110-110	110-110
D3S1282	0.78	170.29	180.8	170.46	168-168	168-168
D3S3523	0.68	170.93	181.87	171.02	188-188	188-188
D3S3723	0.63	171.65	180.08	171.45	161-161	161-161
R1	/	171.8	/	/	271-271	271-271
R2	/	172	/	/	201-201	201-201
R3	/	172.6	/	/	160-160	160-160
D3S1574	0.79	173.08	181.87	174.03	144-144	126-144
D3S3053	0.73	173.15	181.87	/	232-232	236-232
D3S3725	0.84	173.21	181.87	174.44	200-200	198-200
D3S1565	0.63	174.88	186.04	177.94	182-182	182-182
D3S2427	0.85	177.18	188.29	/	230-230	236-230
D3S3730	0.83	179.94	191.79	184.15	274-274	274-274
D3S1262	0.8	187.62	201.14	194.33	114-114	114-114
D3S2436	0.67	188.23	203.28	/	171-171	171-171
D3S1580	0.84	189.94	207.73	202	224-218	218-218
D3S2398	0.77	190.82	209.41	203.31	274-294	294-294
D3S2418	0.71	193.72	215.84	/	106-109	106-109

Table 4.7 Chromosome 3q24-q26 results obtained in patients IV.1 and IV.3 from Family 1 after fine mapping with closely spaced markers. A region of about 17 cM/23 Mb is homozygous in both patients between markers D3S1555 and D3S1574. Note that IV.1 is homozygous for a larger region of about 100 cM (homozygous markers between D3S1271 and D3S1764 have not been included in this table). Blue lines define the minimal critical interval.

Markers Chromosome 3	Family 1 III.1 father	Family 1 III.4 mother	Family 1 IV.1 proband	Family 1 III.2 father	Family 1 III.3 mother	Family 1 IV.2 unaff.	Family 1 IV.3 proband
D3S1764	241-241	241-245	241-241	241-229	241-245	241-145	241-245
D3S1309	167-171	167-165	167-167	171-171	167-165	171-165	171-165
D3S3694	302-294	302-300	302-302	294-292	302-300	294-300	294-300
D3S1569	152-152	152-152	152-152	152-160	152-152	152-152	152-152
D3S1593	226-212	226-212	226-226	212-210	226-212	212-212	212-226
D3S1744	150-146	150-146	150-150	150-146	150-146	150-146	150-150
D3S1306	158-164	158-166	158-158	164-164	158-166	164-166	164-158
D3S1555	211-211	211-209	211-211	203-207	211-209	203-209	203-211
D3S1308	190-190	190-188	190-190	190-194	190-188	190-188	190-190
D3S1299	196-200	196-200	196-196	196-202	196-200	196-200	196-196
D3S1279	266-268	266-276	266-266	266-272	266-276	266-276	266-266
D3S4531	263-251	263-255	263-263	263-263	263-255	263-255	263-263
D3S1746	271-275	271-259	271-271	271-259	271-259	271-259	271-271
D3S1280	210-212	210-208	210-210	210-216	210-208	210-208	210-210
D3S3531	218-218	218-218	218-218	218-220	218-218	218-218	218-218
D3S1607	235-251	235-235	235-235	235-235	235-235	235-235	235-235
D3S3575	188-192	188-196	188-188	188-194	188-196	188-196	188-188
D3S3579	186-172	186-180	186-186	186-184	186-180	186-180	186-186
D3S1268	204-206	204-208	204-204	204-206	204-208	204-208	204-204
D3S3668	228-220	228-222	228-228	228-228	228-222	228-222	228-228
D3S1763	277-273	277-269	277-277	277-273	277-269	277-269	277-277
D3S1614	110-108	110-106	110-110	110-108	110-106	110-106	110-110
D3S1282	168-168	168-166	168-168	168-168	168-166	168-166	168-168
D3S3523	188-188	188-184	188-188	188-188	188-184	188-184	188-188
D3S3723	161-161	161-161	161-161	161-161	161-161	161-161	161-161
R1	271-273	271-275	271-271	271-271	271-275	271-275	271-271
R2	201-201	201-209	201-201	201-201	201-209	201-209	201-201
R3	158-160	160-160	160-160	160-160	160-160	160-160	160-160
D3S1574	144-120	144-136	144-144	144-126	144-136	144-136	126-144
D3S3053			232-232				236-232
D3S3725	200-208	200-210	200-200	200-198	200-210	200-210	198-200
D3S1565	182-182	182-178	182-182	/	182-180	182-180	182-182
D3S2427	230-236	230-230	230-230				236-230
D3S3730	274-284	274-284	274-274	274-274	274-280	274-280	274-274
D3S1262	114-114	114-122	114-114	114-114	114-122	114-122	114-114
D3S2436	171-171	171-175	171-171	183-171	171-175	171-175	171-171
D3S1580	224-228	218-224	224-218	218-218	218-224	218-224	218-218
D3S2398	274-286	294-282	274-294	282-294	294-282		294-294

Table 4.8 Haplotype analysis results across the 3q24-q26 region in Family 1 supporting linkage to this locus. Homozygous results are highlighted in green. Unaffected sib IV.2 shares the same paternal haplotype as affected individual IV.3, but the opposite maternal haplotype. Note that individuals III.4 and III.3 who are sisters, and also the mothers of the two patients IV.1 and IV.3 respectively, have the same haplotype across the region. The region in patient IV.3 is smaller than in IV.1 because of a cross-over occuring in individual III.2 defining the centromeric boundary and a paternal cross-over occuring in patient IV.3 himself which defines the telomeric boundary. Blue lines define the minimal critical interval. Unaff. means individual is unaffected.

4.2.3.2 Family 2

Haplotype analysis across the 3q24-q26 region was performed in Family 2 and was not consistent with linkage, suggesting possible locus heterogeneity (Table 4.9). Fine-mapping was then performed in the two affected individuals from Family 2 (VI.3 and VI.5) filling in gaps in the genome-wide screen with more closely spaced markers (Appendix 1). Four regions were identified in which the two patients were homozygous for the same alleles. Each region was investigated further by genotyping with more markers and testing the parents and unaffected sib to identify a region IBD in this family.

Three consecutive markers were found to be homozygous in both patients on chromosome 2q31-q32 between markers D2S326 and D2S117 (Table 4.10A). The region of shared homozygosity was potentially about 17 cM/22 Mb in size. However, this region was excluded after three more polymorphic markers were chosen and typed in the two affected sibs, their parents and the unaffected sib VI.2 (Table 4.10B).

Three consecutive markers were found to be homozygous in both patients on chromosome 12q21-q22 between markers D12S1052 and D12S95 (Table 4.11A). The region of homozygosity was potentially about 14 cM/17 Mb in size. However, this region was excluded after four more polymorphic markers were chosen and tested in the two affected sibs, and three of them also in the parents and the unaffected sib VI.2 (Table 4.11B).

A third homozygous region in the two affected children was identified on chromosome 18p11 between markers D18S976 and D18S1153 (Table 4.12A). The region was potentially about 22 cM/5 Mb in size. However, this region was excluded after four more polymorphic markers were chosen and tested in the two affected sibs, and three of them also in the parents and the unaffected sib (Table 4.12B). Markers AP001032, AP001793 and AP000864 were found by searching for repeats in these three BAC clones.

The fourth region of interest was located on chromosome 8q24 between markers D8S1128 and D8S373, potentially spanning about 30 cM/15 Mb (Table 4.13). The

Markers Chromosome 3	Family 2 V.5 father	V.6 mother	VI.2 unaff.	VI.3 proband	VI.5 proband
D3S1764				233-237	233-237
D3S1569				164-170	164-164
D3S1744				142-158	142-146
D3S1555	207-209	207-207	207-207	207-207	207-207
D3S1279				264-276	264-268
D3S4531				259-259	255-259
D3S1280	212-216	210-212	210-212	212-212	210-212
D3S3531	214-220	214-216	214-214	214-216	214-214
D3S1607				232-232	232-232
D3S3575	192-192	180-188	180-192	180-192	188-192
D3S3668	224-226	210-228	224-228	224-228	210-224
D3S1763				269-273	269-277
D3S1614				110-110	106-110
D3S1282	166-166	166-168	166-168	166-168	166-166
D3S3523	186-188	186-188	186-188	186-188	188-188
R1				/	272-276
D3S1574	126-128	126-128	126-128	128-128	128-128
D3S3053				224-232	224-232
D3S1565				176-178	178-178
D3S2427				234-238	234-238
D3S2436				167-171	167-171
D3S2398				292-292	288-292

Table 4.9 Haplotype analysis results across the 3q24-q26 region in Family 2. Linkage to this locus was excluded, since a region of homozygosity could not be demonstrated in affected individuals VI.3 and VI.5. Homozygous results are highlighted in green. Not all the markers were typed in the parents and the unaffected sib VI.2. Blue lines indicate the boundaries of the homozygous region in individual IV.3 from Family 1. Unaff. means individual is unaffected.

A)

Markers Chromosome 2	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 VI.3	VI.5
D2S326	0.86	173.06	177.53	/	97-99	97-99
GATA194A05	0.56	176.13	180.79	182.23	235-235	235-235
D2S364	0.80	182.99	186.21	187.67	248-248	248-248
D2S1391	0.79	184.95	186.21	189.15	127-127	127-127
D2S117	0.82	195.58	194.45	194.63	200-204	200-204

B)

Markers Chromosome 2	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 V.5 father	V.6 mother	VI.2 unaff.	VI.3 proband	VI.5 proband
D2S326	0.86	173.06	177.53	/				97-99	97-99
GATA194A05	0.56	176.13	180.79	182.23				235-235	235-235
HOXD10	/	176.94	/	/	113-117	103-101	113-101	113-103	113-103
D2S384	0.72	181.46	185.13	/	257-257	255-259	/	/	257-255
D2S364	0.80	182.99	186.21	187.67				248-248	248-248
D2S1391	0.79	184.95	186.21	189.15				127-127	127-127
D2S152	0.80	188.19	188.11	189.76	228-224	230-239	228-230	228-230	228-230
D2S117	0.82	195.58	194.45	194.63				200-204	200-204

Table 4.10 Genotyping results on chromosome 2q31-q32 in patient VI.3, VI.5 from Family 2 (A) showing a potentially interesting region with three consecutive homozygous markers. The region was excluded by running more polymorphic markers and by testing the parents, who do not share a haplotype in common across this region (B). The two affected sibs are heterozygous for the new markers. Unaff. means idividual is unaffected.

A)

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 VI.3	Family 2 VI.5
D12S1052	0.72	75.32	83.19	87.74	152-156	152-156
D12S326	0.8	77.9	86.4	91.2	216-216	216-216
D12S1708	0.72	83.33	91.46	95.7	166-166	166-166
D12S1064	0.82	90.75	95.03	/	188-188	188-188
D12S95	0.78	92.86	96.09	102.37	172-182	172-182
D12S1044	/	94.1	/	103.94	296-299	296-299

B)

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 V.5 father	V.6 mother	VI.2 unaff.	VI.3 proband	VI.5 proband
D12S1052	0.72	75.32	83.19	87.74				152-156	152-156
D12S1684	0.9	77.19	86.4	89.85	257-255	259-243	255-243	257-259	257-259
D12S326	0.8	77.9	86.4	91.2				216-216	216-216
D12S64	/	79.41	89.42	/	253-253	251-249	253-249	253-251	253-251
D12S1708	0.72	83.33	91.46	95.7				166-166	166-166
D12S1667	0.76	83.66	92.89	97.2	268-278	278-278	278-278	268-278	268-278
D12S1719	0.74	87.09	94.49	98.07				/	242-252
D12S1064	0.82	90.75	95.03	/				188-188	188-188
D12S95	0.78	92.86	96.09	102.37				172-182	172-182
D12S1044	/	94.1	/	103.94				296-299	296-299

Table 4.11 Region on chromosome 12q21-q22 results in patient VI.3, VI.5 from Family 2 (A) showing a potentially interesting region with three consecutive homozygous markers. The region was excluded by running more polymorphic markers and testing the parents and the unaffected sib across this region (B). The two patients are heterozygous for all four new markers. Unaff. means individual is unaffected.

A)

Markers Chromosome 18	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 VI.3	Family 2 VI.5
D18S976	0.86	5.23	12.81	16.6	185-188	180-185
D18S973	/	5.77	/	/	215-215	215-215
D18S452	0.84	5.81	18.7	17.62	125-125	125-125
D18S843	0.78	8.6	28.1	28.11	187-187	187-187
D18S1153	0.78	10.12	35.46	/	138-150	138-150

B)

Markers Chromosome 18	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 V.5 father	V.6 mother	VI.2 unaff.	VI.3 proband	VI.5 proband
D18S976	0.86	5.23	12.81	16.6				185-188	180-185
AP001032	/	5.61	/	/	278-280	278-282	278-278	280-282	278-278
D18S973	/	5.77	/	/				215-215	215-215
D18S452	0.84	5.81	18.7	17.62				125-125	125-125
D18S471	0.66	5.97	18.7	18.37				217-231	217-217
AP001793	/	8.48	/	/	305-295	299-299	305-299	305-299	305-299
D18S843	0.78	8.6	28.1	28.11				187-187	187-187
AP000864	/	8.73	/	/	319-331	317-315	319-315	319-317	319-317
D18S1153	0.78	10.12	35.46	/				138-150	138-150

Table 4.12 Region on chromosome 18p11 results in patient VI.3, VI.5 from Family 2 (A) showing a potentially interesting region with three consecutive homozygous markers. The region was excluded by running more polymorphic markers and testing the parents and the unaffected sib across this region (B). Note that individual VI.2 is homozygous for AP001032 as is his affected brother VI.5. Unaff. means individual is unaffected.

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 VI.3	VI.5
D8S1128	0.76	128.26	139.53	135.57	/	238-242
D8S284	0.83	131.18	143.82	139.79	296-296	296-296
D8S1462	0.75	135.06	/	/	160-160	160-160
D8S1100	0.65	136.77	154.02	148.81	182-182	182-182
D8S272	0.8	137.41	154.02	/	253-253	253-253
D8S373	0.78	144.00	167.9	166.08	264-264	264-276

Table 4.13 Genotyping results on chromosome 8q24 obtained in patients VI.3 and VI.5 from Family 2 after the second-pass screen, showing a region of shared homozygosity in the two affected individuals between markers D8S1128 and D8S373.

region was investigated further and more closely spaced markers, approximately 1-3 cM apart, were used to genotype the two affected children, their parents and two unaffected sibs (VI.2 and VI.4). Details of markers are shown in Table 2.6 and their positions in Table 4.14. In Table 4.14, different colours represent different haplotypes to show how they segregate in the family. Some markers were uninformative or semi-informative. In this case, the allele sizes were left in black for the children. DNA from unaffected sib VI.4 was obtained later in the study, therefore not all the markers were typed. Haplotype analysis across the region in the parents and the two unaffected sibs was consistent with linkage to this locus (Table 4.14). Both affected sibs continued to show homozygosity over the whole region. The parents shared a haplotype in common which was inherited by the two affected children. Both unaffected sibs were heterozygous for all the markers typed, except for the uninformative or semi-informative markers. The shared homozygous region of 13 cM/ 8 Mb was defined between markers AC103725 and AC009682 at the centromeric end and between AC087711 and AC087711-2 at the telomeric end. A recombination event that occurred in patient VI.5 between AC087711 and AC087711-2 defines the telomeric boundary of the region (Table 4.14).

Haplotype analysis across the 8q24 region was performed in affected children IV.1 and IV.3 from Family 1 using five polymorphic markers and was not consistent with linkage (Table 4.15).

4.2.3.3 Family 4

An independent genome-wide screen was performed in affected individual II.6 from consanguineous Family 4. The haplotype of this patient was not consistent with linkage to any of the putative candidate JATD loci identified or to the *EVC/EVC2* locus. Furthermore, in this family there were six unaffected sibs who were thought to be useful to support the hypothesis for linkage to any homozygous region eventually identified in the affected child.

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	Family 2 V.5 father	V.6 mother	VI.2 unaff.	VI.4 unaff.	VI.3 proband	VI.5 proband
D8S1128	0.76	128.26	139.53	135.57						238-242
D8S1720	0.82	128.62	140.61	137.06	156-152	158-158	156-158		156-158	156-158
D8S1732	0.78	130.04	143.29	139.58					212-214	212-214
D8S1701	0.7	130.56	142.22	/	275-279	/	279-279		275-279	275-279
AC131568	/	130.65	/	/	182-188	176-192	182-192		182-176	182-176
AC103725	/	130.758	/	/	219-225	223-233	219-233		219-223	219-223
AC009682	/	130.952	/	/	239-237	239-235	239-235		239-239	239-239
D8S284	0.83	131.18	143.82	139.79					296-296	296-296
D8S1712	0.72	131.323	144.36	/	225-225	225-223	225-223		225-225	225-225
D8S557	0.76	132.761	145.97	142.18	202-212	202-216	202-216	212-216	202-202	202-202
D8S558	0.86	133.163	145.97	143.14	171-157	171-181	171-181	157-181	171-171	171-171
D8S529	0.81	133.770	148.12	144.75					235-235	235-235
D8S256	0.83	134.115	148.12	145.26	232-232	232-226	232-226		232-232	232-232
D8S1462	0.75	135.06	/	/	160-164	160-164	160-164	164-164	160-160	160-160
D8S1710	0.78	135.32	150.8	146.68	255-259	255-255	255-255		255-255	255-255
D8S537	0.89	135.51	150.8	147.41	180-180	180-182	180-182		180-180	180-180
D8S1100	0.65	136.77	154.02	148.81	182-188	182-188	182-188		182-182	182-182
D8S1761	0.71	137.25	154.02	150.07	290-290	290-290	290-290		290-290	290-290
D8S272	0.8	137.41	154.02	/	253-249	253-257	253-257	249-257	253-253	253-253
D8S1837	0.81	138.91	156.59	152.44	233-231	233-247	233-247	231-247	233-233	233-233
AFM316xb1	/	138.933	/	/	193-191	193-191	193-191	191-191	193-193	193-193
AC087711	/	138.939	/	/	204-200	204-198	204-198	200-198	204-204	204-204
AC087711-2	/	138.955	/	/	248-250	248-250	248-250		248-248	248-250
AC068467	/	139.150	/	/	209-211	209-211	209-211		209-209	209-211
AC105130-R1	/	139.350	/	/	257-245	257-245	257-245		257-257	257-245
AC053480-R1	/	140.000	/	/	285-287	285-287	285-287		285-285	285-287
D8S1743	0.83	140.33	162.94	/	220-226	220-246	220-246	226-246	220-220	220-246
D8S373	/	145.74	167.9	166.08					264-264	264-276
D8S1926	/	145.74	167.9	166.08					246-246	246-248

Table 4.14 Haplotype analysis results across the 8q24 region in Family 2 consistent with linkage to this locus. The two affected children share a homozygous region between markers AC103725 and AC87711-2. Pink lines define the minimal critical interval.

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	IV.3
D8S284	0.83	131.18	143.82	139.79	278-296	270-296
D8S558	0.86	133.163	145.97	143.14	163-169	159-169
D8S1462	0.75	135.06	/	/	156-160	156-156
D8S1100	0.65	136.77	154.02	148.81	185-188	188-188
D8S1837	0.81	138.91	156.59	152.44	243-243	241-243

Table 4.15 Genotyping results across the 8q24 region obtained in patients IV.1 and IV.3 from Family 1, showing no evidence of linkage to this locus.

Following the linkage screen in affected individual II.6, four regions were identified in which he was homozygous for consecutive markers. Each region was investigated further by genotyping with more closely spaced markers and testing the parents and the six unaffected sibs to identify a region IBD in the affected child.

Three consecutive markers were found to be homozygous on chromosome 1q24-q26 between markers D1S2851 and D1S1660 (see Appendix 1). The region of homozygosity was potentially about 22 cM/28 Mb in size. However, this region was excluded after seven further polymorphic markers were chosen and typed in the affected individual, and five of them also in the parents (I.1 and I.2) and six unaffected sibs (Table 4.16).

A second homozygous region in the affected child was identified on chromosome 3q28-q29 from marker D3S2398 to D3S1311 at the telomeric end (see Appendix 1). The region was potentially about 18 cM/8 Mb in size. However, this region was excluded after five more polymorphic markers were chosen and typed in the affected child, his parents and the six unaffected (Table 4.17). The patient II.6 showed homozygosity across the region, but so did two of his unaffected sibs (II.2 and II.7).

Three consecutive markers on chromosome 14q21 were also homozygous in patient II.6 from D14S587 to D14S592 (Table 4.18). Further typing was carried out in the whole family and almost all the markers were uninformative. DNA from a second affected baby girl (II.8) was later obtained and she was found to be heterozygous for two markers. However results were inconclusive and more typing needs to be done to completely exclude this region. The interval between the two markers heterozygous in II.8 is in fact quite large (about 7 cM).

A fourth homozygous region was identified on chromosome 14q24-q32. Four consecutive markers were homozygous in proband II.6 between D14S74 and D14S81 spanning about 19 cM/14 Mb (Table 4.19). Further polymorphic markers were typed in the whole family and a few of them also in the affected girl II.8. Both affected children II.6 and II.8 showed homozygosity throughout the region, although most of the markers were uninformative in the father. Therefore, unaffected individuals II.2 and II.3 were also homozygous across the interval. Unaffected children II.2 and II.3 could have

Family 4

Markers Chrom. 1	Het	Mb Build 33	cM Marsh.	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 unaff.	II.3 unaff.	II.4 unaff.	II.5 unaff.	II.7 unaff.	II.6 proband
D1S2851	**0.83**	**167.01**	**188.32**	**172.9**									270-274
D1S1589	**0.77**	**170.84**	**192.05**	**176.25**									**205-205**
D1S2786	**0.72**	**173.87**	**193.76**	**178.47**	124-116	116-126	124-126	116-116	124-126	124-126	/	124-116	124-116
D1S2751	**0.71**	**176.53**	**194.89**	**180.36**	286-284	286-292	286-292	284-286	286-292	286-292	286-292	286-286	286-286
D1S466	**0.77**	**178.82**	**198.3**	**183.53**									**230-230**
D1S191	**0.73**	**182.34**	**200.96**	**186.76**	247-245	243-245	247-245	245-243	247-245	247-245	247-245	247-243	247-243
D1S518	**0.84**	**184.07**	**202.19**	**188.02**	206-214	206-198	206-198	214-206	206-198	206-198	206-198	206-206	**206-206**
D1S238	**0.86**	**184.67**	**202.73**	**188.55**	190-192	/	190-194	192-190	190-194	190-194	/	/	/
D1S3468	**0.81**	**188.01**	**205.40**	/									187-187
D1S533	**0.81**	**190.52**	**209.15**	**193.04**	249-237	245-245							249-245

Table 4.16 Haplotype analysis results on chromosome 1q in consanguineous Family 4 are not consistent with linkage to this locus. Heterozygosity, physical and genetic distances of each marker are indicated. Homozygous results in affected child II.6 obtained following the genome-wide screen are shown in bold black. The homozygous region was interrupted by heterozygous markers D1S2786 and D15S191. Moreover, D1S2751 and D1S518 were also homozygous in unaffected sib II.7. Markers which could not be typed for a particular individual are marked with a bar. Unaff. means individual is unaffected.

Family 4

Markers Chrom. 3	Het	Mb Build 33	cM Marsh.	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 unaff.	II.3 unaff.	II.4 unaff.	II.5 unaff.	II.7 unaff.	II.6 proband
D3S1580	0.84	189.95	207.73	202									224-228
D3S3530	0.80	190.56	209.41	/	143-137	143-147		143-143				/	143-143
D3S2398	0.77	190.82	209.41	203.31	290-274	290-298	290-298	290-290	/	274-298	274-298	290-290	290-290
D3S2747	0.75	191.68	212.61	204.73	283-289	283-287	283-287	283-283	289-283	289-287	289-287	283-283	283-283
D3S3054	0.69	191.95	214.45	/	391-383	391-383		391-391				391-391	391-391
D3S1601	0.86	193.08	214.45	208.14	216-210	216-206	216-206	216-216	210-216	210-206	210-206	216-216	216-216
D3S2418	0.71	193.72	215.84	/	94-94	94-103	94-103	94-94	94-94	94-103	94-103	94-94	94-94
D3S2748	0.73	195.13	217.24	/	260-272	260-260	260-260	260-260	272-260	272-260	272-260	260-260	260-260
D3S1311	0.82	198.42	224.88	220.19	133-145	133-147	133-147	133-133	145-133	145-147	145-147	133-133	133-133

Table 4.17 Haplotype analysis results on chromosome 3q in consanguineous Family 4 are not consistent with linkage to this locus. Heterozygosity, physical and genetic distances of each marker are indicated. Unaffected children II.2 and II.7 are also homozygous across the same region as proband II.6. Almost all markers are informative. Markers which could not be typed for a particular individual are marked with a bar. Unaff. means individual is unaffected.

Family 4

Markers Chrom. 14	Het	Mb Build 33	cM Marsh.	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 unaff.	II.3 unaff.	II.4 unaff.	II.5 unaff.	II.7 unaff.	II.6 proband	II.8 proband
D14S583		42.21	/	/									308-312	
D14S978	0.84	49.90	53.19	/	256-270	256-276	256-276	256-256	270-256	256-256	256-256	256-256	256-256	
D14S587	0.83	52.36	55.82	/	246-262	246-250	246-250	246-246	262-246	246-246	246-246	246-246	246-246	262-250
D14S276	0.75	53.67	56.36	/	198-196	198-198	198-198	198-198	196-198	198-198	198-198	198-198	198-198	
D14S285	0.82	54.96	59.43	56.52	251-251	251-253	251-253	251-251	251-251	251-251	251-251	251-251	251-251	
D14S980	0.86	55.14	60.43	56.77	206-206	206-210	206-210	206-206	206-206	206-206	206-206	206-206	206-206	
D14S607	/	55.61	/	/	266-266	266-268	266-268	266-266	266-266	266-266	266-266	266-266	266-266	
D14S274	0.72	55.65	63.25	58.15	264-264	264-270	264-270	264-264	264-264	264-264	264-264	264-264	264-264	264-270
D14S1038	0.79	57.61	66.81	60.06	212-212	212-212	212-212	212-212	212-212	212-212	212-212	212-212	212-212	
D14S592	0.68	59.39	66.81	60.5	253-253	253-253	253-253	253-253	253-253	253-253	253-253	253-253	253-253	
D14S63	0.77	62.64	69.18	63.50									174-186	

Table 4.18 Haplotype analysis results on chromosome 14q in consanguineous Family 4. Heterozygosity, physical and genetic distances of each marker are indicated. Almost all markers are semi- or uninformative. Therefore, the unaffected sibs are also homozygous across the region as well as affected child II.6. However, proband II.8 is heterozygous for two markers in the region. Unaff. means individual is unaffected.

Family 4

Markers Chrom. 14	Het	Mb Build 33	cM Marsh.	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 unaff.	II.3 unaff.	II.4 unaff.	II.5 unaff.	II.7 unaff.	II.6 proband	II.8 proband
D14S74	0.79	76.65	87.36	78.19									301-303	
D14S739	0.83	80.26	91.62	81.73	192-176	192-168	192-168	176-192	176-192	192-168	192-168	192-168	192-192	192-192
D14S616	0.7	83.18	92.69	83.38	219-231	219-219	219-219	231-219	231-219	219-219	219-219	219-219	219-219	
D14S1052	0.74	84.40	93.76	84.31	214-214	214-214	214-214	214-214	214-214	214-214	214-214	214-214	214-214	
D14S67	0.86	86.38	95.89	/	264-264	264-260	264-260	264-264	264-264	264-260	264-260	264-260	264-264	
D14S68	0.91	86.62	95.89	/	137-137	137-129	137-129	137-137	137-137	137-129	137-129	137-129	137-137	137-137
D14S256	0.71	87.20	96.42	87.01	141-141	141-129	141-129	141-141	141-141	141-129	141-129	141-129	141-141	
D14S1005	0.71	87.37	96.42	87.25	259-259	259-257	259-257	259-259	259-259	259-257	259-257	259-257	259-259	
D14S1044	0.67	88.06	99.88	90.62	199-199	199-203	199-203	199-199	199-199	199-203	199-203	199-203	199-199	199-199
D14S995	0.82	89.57	105	/	241-241	241-241	241-241	241-241	241-241	241-241	241-241	241-241	241-241	
D14S280	0.68	90.17	105	94.42	256-256	256-264	256-264	256-256	256-256	256-264	256-264	256-264	256-256	
D14S617	0.78	90.19	105.53	94.47	146-146	146-142	146-142	146-146	146-146	146-142	146-142	146-142	146-146	146-146
D14S977	0.82	91.16	107.13	/	224-220	220-218	224-218	220-220	220-220	224-218	224-218	220-218	224-220	
D14S81	0.82	91.77	108.22	96.87									282-286	

Table 4.19 Haplotype analysis results on chromosome 14q24-q32 in consanguineous Family 4. Heterozygosity, physical and genetic distances of each marker are indicated. Almost all markers are semi- or uninformative. Therefore, unaffected sibs II.2 and II.3 are also homozygous across the region as well as affected children II.6 and II.8. However, II.2 and II.3 might have inherited a different paternal haplotype. Unaff. means individual is unaffected.

inherited a different paternal haplotype compared to their affected sibs and more typing needs to be carried out in this family to corroborate these results. Nevertheless, following the genome-wide screen, this locus was thought to represent the most interesting region.

4.2.4 Summary

A genome-wide screen was carried out in five affected children from Families 1, 2 and 4. A region which was IBD in all 5 patients was not identified, not even after a second-pass screen with more closely spaced markers was performed. Regions found to be homozygous in patients from each individual family were then studied further.

A region IBD of 17 cM/23 Mb was identified in Family 1 on chromosome 3q24-q26. The region extends as centromeric to D3S1555 and as telomeric to D3S1574 as these were two heterozygous markers in affected individual IV.3. Linkage to this locus was supported by the haplotype results in the whole family. Family 2 did not show evidence for linkage to this locus implying the existence of a second disease locus in Family 2 and therefore locus heterogeneity.

In Family 2, an IBD region of 13 cM/8 Mb was subsequently identified on chromosome 8q24 between markers AC103725 and AC087711-2, which defined the centromeric and telomeric boundaries respectively. Evidence for linkage to this locus was supported by the haplotype results in the whole family. Linkage to this locus could not be demonstrated in Family 1.

In order to confirm whether or not the two identified loci were linked to the JATD locus, a genotyping screen was performed in all the other JATD consanguineous families looking for homozygosity in the two regions as well as genotyping four non-consanguineous families to see if markers across the regions segregate in a way that was consistent with linkage. In the final stages of the project, a homozygous region of 19 cM/14 Mb on chromosome 14q24-q32 was identified in Family 4 and genotyping across this locus was performed only in non-consanguineous Family 15.

4.2.5 Linkage screens carried out by collaborators

Two independent genome-wide screens were carried out by our collaborators Dr Colin Johnson in Birmingham and Dr Valerie Cormier-Daire in Paris, using the homozygosity mapping approach in affected children from their own collection of consanguineous families. Results were shared between the three groups and it became clear that JATD was genetically heterogeneous.

A locus on chromosome 12q23-q24 between markers D12S1342 and D12S1339 was mapped by the Paris group supported by data from two consanguineous families.

The Birmingham group identified a locus on chromosome 15q13 supported by data from five consanguineous families with a maximum two-point lod score of $Z = 3.77$ at $\theta = 0$ obtained at D15S1031. The minimal critical interval was 1.2 cM (~1.5 Mb) in size, and contained 16 known or predicted genes. Interestingly, *FORMIN-1* and *GREMLIN* lie in this interval and initially appeared to be good candidates for JATD but none of the linked families were found to harbour mutations in the coding sequence of either gene (Morgan *et al.*, 2003). These results were published and a copy of the paper is included at the back of this thesis.

As part of the collaboration, it was decided that each group should type their own collection of families across all four regions. For this purpose, closely spaced polymorphic markers were shared between the three groups. In case one or more families were found to be linked to a locus discovered by another group, it was agreed that DNA samples should then be shared and each group continue to work on their own region.

4.3 Genotyping analysis of JATD families

Once the two candidate regions on chromosome 3q24-q26 and 8q24 were identified in Families 1 and 2, respectively, a genotyping screen was performed in 10 other JATD consanguineous families looking for homozygosity in the affected individuals across the two loci, as well as in four non-consanguineous families to see if markers across the

regions segregate in a way that was consistent with linkage. The aim of the screen was to investigate whether these two regions were linked to the JATD disease locus by looking for evidence of linkage in other families, as well as to narrow down the candidate regions. In this study, the same families were also genotyped across the 12q23-q24 and the 15q13 regions identified by our collaborators in Birmingham and Paris. DNA samples from Families 8, 9 and 16 were provided by Dr Colin Johnson and DNA samples from Family 10 were sent by Dr Valerie Cormier-Daire for further genotyping. Genotyping results shown in this chapter in the above four families were obtained by myself. Extra markers were ascertained in an attempt to further narrow down the candidate regions prior to searching for candidate genes.

4.3.1 Chromosome 3q24-q26 locus

Consanguineous Families 3-12 and non-consanguineous Families 13-16 were genotyped across the 3q24-q26 locus with a selection of evenly spaced markers across the region homozygous in individual IV.3 from Family 1 (see Table 4.7). Not all the markers used for typing Family 1 were used in the other families. Details of all the markers are shown in Table 2.5. Pedigrees of the families are shown in Chapter 2.

4.3.1.1 Genotyping in JATD families consistent with linkage to 3q24-q26

Consanguineous Families 5, 6 and 8 supported the hypothesis for linkage to 3q24-q26 with affected children demonstrating regions of homozygosity. Non-consanguineous Family 14 was also thought to show evidence of linkage to this locus as the two affected children shared the same pair of haplotypes across the region, which was different in the unaffected sib.

In Family 5, patient II.1 demonstrated a region of homozygosity between markers D3S1306 and D3S1574, which, assuming linkage, did not help to narrow down the interval defined in patient IV.3 from Family 1 (Table 4.20). DNA from the mother was

Family 5

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	II.1 proband
D3S1306	0.72	149.08	164.25	154.18	163-163	163-163
D3S1308	0.69	150.31	165.85	157.07	193-189	193-193
D3S1279	0.85	152.3	169.6	160.19	264-268	264-264
D3S4531	0.69	152.53	169.6	160.46	261-265	261-261
D3S1746	0.85	153.01	169.6	/	275-259	275-275
D3S1280	0.75	153.38	170.14	161.6	212-210	212-212
D3S3531	0.73	155.49	170.14	162.79	254-254	254-254
D3S1607	0.80	158.24	172.27	164.61	235-235	/
D3S3575	0.79	160.63	173.34	165.61	196-190	196-196
D3S3668	0.83	165.72	175.47	167.47	228-224	228-228
D3S1614	0.83	169.61	177.75	169.98	116-106	116-116
D3S1282	0.78	170.29	180.8	170.46	168-172	168-168
D3S3523	0.68	170.93	181.87	171.02	188-186	188-188
D3S1574	0.79	173.08	181.87	174.03	137-137	137-137

Table 4.20 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 5. Patient II.1 displays homozygosity across the region homozygous in patient IV.3 from Family 1 (defined by blue lines). DNA from the mother was not available for genotyping.

not available for typing. DNA stock from affected child II.1 was of limited quantity and of poor quality, therefore very few markers were used for genotyping. Attempts to amplify the patient DNA using GenomiPhi™ DNA Amplification Kit (Amersham Biosciences) failed. This kit was chosen because it allows DNA amplification of small quantities of starting material as detailed in section 2.3.1.4. Although this technique allows amplification of good quality DNA, poor quality DNA, as in the case of this patient, does not amplify well.

In Family 6, affected individual II.2 demonstrated homozygosity across the candidate locus, from D3S1744 to D3S3725 (Table 4.21). Linkage was supported by the haplotype in the family. Unaffected individual II.1 was heterozygous for most of the markers typed as he had inherited different haplotypes from both parents compared to his affected sister II.2. Assuming linkage to this locus, the homozygous region in the patient did not help to narrow down the region of homozygosity identified in patient IV.3 from Family 1.

In Family 8, a region of homozygosity was demonstrated between markers D3S1593 and D3S3523 in patient II.1 (Table 4.22). The region of about 19 cM/25 Mb did not extend as telomeric to D3S1574 as in affected individual IV.3 from Family 1 and assuming linkage, the candidate region was slightly narrowed down at the telomeric end (Table 4.22). Markers D3S3705, D3S1237, D3S3710 and D3S3692 were ascertained and typed only in this family to demonstrate a convincing block of homozygosity in patient II.1.

Non-consanguineous Family 14 was genotyped, as there were two affected children (II.3 and II.4) and one unaffected sib (II.1). DNA sample from the latter (II.1) was obtained later in the course of the project. A shared pair of haplotypes was demonstrated in the two affected individuals across the entire region and was consistent with linkage to the 3q24-q26 locus (Table 4.23). Genotyping was then performed in the unaffected sib II.1 and she was found to share the same pair of haplotypes as her affected sibs up to marker D3S3580, then a maternal cross-over occurred and she inherited the different maternal haplotype compared to her affected sibs. Therefore, assuming linkage in this family, the candidate region was narrowed down at the centromeric end between markers D3S3575 and D3S3580 due to the recombination

Family 6

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 proband
D3S1744	0.77	148.37	161.04	153.36	150-142	150-150	142-150	150-150
D3S1306	0.72	149.08	164.25	154.18	157-163	157-163	163-163	157-157
D3S1308	0.69	150.31	165.85	157.07	193-193	193-191	193-191	193-193
D3S1279	0.85	152.3	169.6	160.19	269-271	269-267	271-267	269-269
D3S1746	0.85	153.01	169.6	/	267-267	267-267	267-267	267-267
D3S1607	0.8	158.24	172.27	164.61	246-236	246-246	236-246	246-246
D3S3575	0.79	160.63	173.34	165.61	196-192	196-192	192-192	196-196
D3S3579	0.79	161.86	173.87	166.45	180-184	180-186	184-186	180-180
D3S1268	0.86	165.29	174.94	/	194-206	194-206	206-206	194-194
D3S3668	0.83	165.72	175.47	167.47	222-228	222-224	228-224	222-222
D3S1763	0.8	168.64	176.54	168.52	276-276	276-272	276-272	276-276
D3S1614	0.83	169.61	177.75	169.98	110-110	110-110	110-110	110-110
D3S1282	0.78	170.29	180.8	170.46	166-176	166-166	176-166	166-166
D3S3523	0.68	170.93	181.87	171.02	191-189	191-187	189-187	191-191
D3S1574	0.79	173.08	181.87	174.03	126-136	126-130	136-130	126-126
D3S3725	0.84	173.21	181.87	174.44	197-197	197-203	197-203	197-197

Table 4.21 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 6. Patient II.2 displays homozygosity across the region homozygous in patient IV.3 from Family 1 (defined by blue lines). Unaffected sib II.1 has inherited different haplotypes from both parents compared to his affected sister II.2. Unaff. means individual is unaffected.

Family 8

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D3S1764	0.79	140.46	152.62	145.53	232-248	240-252	232-240
D3S3694	0.84	143.47	153.74	148.82	306-300	302-300	306-302
D3S1593	0.78	146.6	161.04	152.31	212-226	226-226	212-226
D3S1744	0.77	148.37	161.04	153.36	161-153	161-137	161-161
D3S1306	0.72	149.08	164.25	154.18	164-164	164-164	164-164
D3S1308	0.69	150.31	165.85	157.07	193-191	193-193	193-193
D3S3705	0.67	150.32	165.32	157.53	227-225	227-227	227-227
D3S1299	0.62	151.46	166.93	158.56	200-198	200-188	200-200
D3S1279	0.85	152.3	169.6	160.19	265-263	265-267	265-265
D3S4531	0.69	152.53	169.6	160.46	260-256	260-260	260-260
D3S1746	0.85	153.01	169.6	/	251-247	251-251	251-251
D3S1280	0.75	153.38	170.14	161.6	216-212	216-216	216-216
D3S1237	0.79	153.44	172.27	161.34	317-317	317-315	317-317
D3S3710	0.74	155.38	170.67	/	246-252	246-248	246-246
D3S3531	0.73	155.49	170.14	162.79	218-218	218-218	218-218
D3S1607	0.8	158.24	172.27	164.61	248-246	248-236	248-248
D3S3692	0.62	159.89	172.27	165.07	241-253	241-241	241-241
D3S3575	0.79	160.63	173.34	165.61	193-193	193-189	193-193
D3S3579	0.79	161.86	173.87	166.45	186-182	186-184	186-186
D3S1268	0.86	165.29	174.94	/	204-204	204-196	204-204
D3S3668	0.83	165.72	175.47	167.47	228-226	228-228	228-228
D3S1763	0.8	168.64	176.54	168.52	276-276	276-276	276-276
D3S1614	0.83	169.61	177.75	169.98	110-106	110-108	110-110
D3S1282	0.78	170.29	180.8	170.46	170-170	170-166	170-170
D3S3523	0.68	170.93	181.87	171.02	185-189	189-187	185-189
D3S3723	0.63	171.65	180.08	171.45	165-161	163-161	165-163
R3	/	171.8	/	/			
D3S1574	/	172	/	/			

Table 4.22 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 8. In patient II.1 there is a region of homozygosity between markers D3S1593 and D3S3523. Note how some markers are semi-informative as they are homozygous in at least one of the two parents as well as in the affected child. Blue lines define the homozygous region in patient IV.3 from Family 1. The telomeric boundary is now between D3S1282 and D3S3523 (red line) narrowing down the candidate region by about 3 cM/3 Mb.

Family 14

Markers Chromosome 3	I.1 father	I.2 mother	II.1 unaff.	II.3 proband	II.4 proband	
D3S1306	158-164	164-164	158-164	158-164	158-164	
D3S1308	193-195	193-191	193-193	193-193	193-193	informative
D3S1279	265-269	267-267		265-267	265-267	
D3S4531	263-259	259-251	263-259	263-259	263-259	informative
D3S1280	212-212	214-214			212-214	
D3S1607	248-236	244-232	248-244	248-244	248-244	informative
D3S3575	188-192	190-196	188-190	188-190	188-190	informative
D3S3580	242-240	242-244	242-244	242-242	242-242	**cross-over in II.1**
AC026118	187-181	181-187	187-187	187-181	187-181	
AC024221	237-239	237-237	237-237	237-237	237-237	
AC069224	221-219	223-223	221-223	221-223	221-223	
D3S3579	184-186	186-184	184-184	184-186	184-186	
D3S1268	206-208	196-208	206-208	206-196	206-196	informative
D3S3668	220-228	224-224	220-224	220-224	220-224	
D3S1763	268-268	264-260	268-260	268-264	268-264	
D3S1614	106-104	112-116	106-116	106-112	106-112	informative
D3S1282	175-167	167-167	175-167	175-167	175-167	
D3S3523	186-186	184-190	186-190	186-184	186-184	
D3S3723	163-163	161-161		163-161	163-161	
R1	271-265	267-271	271-271	271-267	271-267	informative
R2	203-211	201-201		203-201	203-201	
R3	288-258	288-288		288-288	288-288	
D3S1574	127-137	127-129	127-127	127-127	137-127	**cross-over in II.1**
D3S3053	234-226	234-226	234-234	234-234	226-234	**and II.4**
D3S3725	199-203	209-209	199-209	199-209	203-209	
D3S1565	176-180	182-178	176-182	176-182	180-182	informative

Table 4.23 Haplotype analysis results of chromosome 3q24-q26 region in non-consanguineous Family 14. Patients II.3 and II.4 display the same haplotypes across the region from D3S1306 to R3 (highlighted in blue box) where a paternal cross-over occurs in patient II.4. Unaffected sib II.1 shares the same haplotypes as her affected sibs up to D3S3580 (upper red line). A maternal cross-over in II.1 allows the region to be narrowed down at the centromeric end between D3S3575 and D3S3580. A second maternal cross-over in II.1 occurs at the telomeric end between R1 and D3S1574. Blue lines define the homozygous interval in Family 1. The lower red line indicates the telomeric end of the region homozygous in patient II.1 from Family 8. Fully informative markers are indicated. The yellow box shows the region in which unaffected individual II.1 has a different haplotype from her sibs.

event which occurred in individual II.1 (Table 4.23). Different colours for allele sizes represent different haplotypes to show how they segregate in the family. Some markers were uninformative or semi-informative. In this case, the allele sizes were left in black for the children. Markers D3S3580 as well as repeats identified in BAC clones AC026118, AC024221 and AC069224 were ascertained and typed only in this family in the attempt to better refine the cross-over in individual II.1.

A summary of the genotyping for affected individuals in families 1, 5, 6, 8 and for the unaffected sib in Family 14 is given in Table 4.24. The common region of homozygosity of 10.3 Mb and 5.4 cM, according to the deCode genetic map, was defined between markers D3S3575 and D3S3523. The recombination event which occurred in unaffected individual II.1 in non-consanguineous Family 14 defines the centromeric boundary. The telomeric boundary is defined by the first heterozygous markers in patient II.1 from Family 8.

It was noted that for three consecutive markers in the interval certain alleles appeared more often than others amongst the affected individuals of the same ethnic origin (Table 4.24). Allele 186 for marker D3S3579 appeared in Pakistani patients IV.1 and IV.3 (Family 1), II.1 (Family 8), II.3 and II.4 (Family 14), but not in the unaffected sib II.1 of Family 14 (Table 5.19) or in affected Turkish child II.2 from Family 6. Allele 204 for marker D3S1268 was present in Pakistani patients IV.1, IV.3 (Family 1) and II.1 (Family 8). Allele 228 for marker D3S3668 was detected in Pakistani affected individuals IV.1 and IV.3 (Family 1), II.1 (Family 5) and II.1 (Family 8). This suggested the presence of a shared ancestral chromosomal segment in the Pakistani population inherited from a distant common ancestor and was consistent with the possibility of linkage to this locus.

Consanguineous Family 7 could be consistent with linkage to this locus, but no DNA from the affected, deceased child (II.3) was available. Haplotype analysis was performed in the parents (I.1 and I.2) and unaffected sibs (II.1 and II.4) and was consistent with linkage (Table 4.25). The two unaffected sibs were heterozygous across the region and the parents shared a haplotype in common, so the affected child could, in principle, be homozygous. DNA extraction from paraffin embedded tissues from the proband II.3 was performed using the QIAamp tissue kit (QIAGEN Ltd.), but with no

Markers Chromosome 3	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1 Pakistani	Family 1 IV.3 Pakistani	Family 5 II.1 Pakistani	Family 6 II.2 Turkish	Family 8 II.1 Pakistani	Family 14 II.1 Pakistani
D3S1764	140.46	152.62	145.53	241-241	241-245				
D3S1309	142	153.74	146.63	167-167	171-165				
D3S3694	143.47	153.74	148.82	302-302	294-300				
D3S1569	144.65	158.38	150.58	152-152	152-152				
D3S1593	146.6	161.04	152.31	226-226	212-226				
D3S1744	148.37	161.04	153.36	150-150	150-150		150-150	161-161	
D3S1306	149.08	164.25	154.18	158-158	164-158	163-163	157-157	164-164	158-164
D3S1555	150.08	165.32	156.59	211-211	203-211				
D3S1308	150.31	165.85	157.07	190-190	190-190	193-193	193-193	193-193	193-193
D3S3705	150.32	165.32	157.03					227-227	
D3S1299	151.46	166.93	158.56	196-196	196-196			200-200	
D3S1279	152.3	169.6	160.19	266-266	266-266	264-264	269-269	265-265	
D3S4531	152.53	169.6	160.46	263-263	263-263	261-261		260-260	263-259
D3S1746	153.01	169.6	/	271-271	271-271	275-275	267-267	251-251	
D3S1280	153.38	170.14	161.6	210-210	210-210	212-212		216-216	
D3S1237	153.44	172.27	161.34					317-317	
D3S3710	155.38	170.67	/					246-246	
D3S3531	155.49	170.14	162.79	218-218	218-218	254-254		218-218	
D3S1607	158.24	172.27	164.61	235-235	235-235		246-246	248-248	248-244
D3S3692	159.87	172.27	165.07					241-241	
D3S3575	160.63	173.34	165.61	188-188	188-188	196-196	196-196	193-193	188-190
D3S3580	160.900	173.34	166.11						242-244
AC026118	160.994	/	/						187-187
AC024221	161.424	/	/						237-237
AC069224	161.800	/	/						221-223
D3S3579	161.86	173.87	166.45	186-186	186-186		180-180	186-186	184-184
D3S1268	165.29	174.94	/	204-204	204-204		194-194	204-204	206-208
D3S3668	165.72	175.47	167.47	228-228	228-228	228-228	222-222	228-228	220-224
D3S1763	168.64	176.54	168.52	277-277	277-277		276-276	276-276	268-260
D3S1614	169.61	177.75	169.98	110-110	110-110	116-116	110-110	110-110	106-116
D3S1282	170.29	180.8	170.46	168-168	168-168	168-168	166-166	170-170	175-167
D3S3523	170.93	181.87	171.02	188-188	188-188	188-188	191-191	185-189	186-190
D3S3723	171.65	180.08	171.45	161-161	161-161			163-165	
R1	171.8	/	/	271-271	271-271				271-271
R2	172	/	/	201-201	201-201				
R3	172.6	/	/	160-160	160-160				
D3S1574	173.08	181.87	174.03	144-144	126-144	137-137	126-126		127-127
D3S3053	173.15	181.87	/	232-232	236-232				234-234
D3S3725	173.21	181.87	174.44	200-200	198-200		197-197		199-209
D3S1565	174.88	186.04	177.94	182-182	182-182				176-182
D3S2427	177.18	188.29	/	230-230	236-230				
D3S3730	179.94	191.79	184.15	274-274	274-274				
D3S1262	187.62	201.14	194.33	114-114	114-114				
D3S2436	188.23	203.28	/	171-171	171-171				
D3S1580	189.94	207.73	202	224-218	218-218				
D3S2398	190.82	209.41	203.31	274-294	294-294				
D3S2418	193.72	215.84	/	106-109	106-109				

Table 4.24 Summary diagram of the genotyping in affected individuals in Families 1, 5, 6 and 8 as well as unaffected individual in Family 14. These data are consistent with linkage to 3q24-q26. Homozygous typings are highlighted in bright green. The region in which unaffected individual II.1 (Family 14) does not share the same haplotypes as her two affected sibs is highlighted in yellow. Untyped markers are highlighted in white. The markers highlighted in pale blue between the two red lines denote the region of shared homozygosity flanked by markers D3S3575 and D3S3523. Note how the genetic distances of some markers (highlighted in orange) sometimes differ from their positions in Mb making the integration between the two maps often difficult. The ethnic origins of the families are also shown.

Family 7

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.4 unaff.
D3S1306	0.72	149.08	164.25	154.18	157-157	157-161	157-157	157-161
D3S1308	0.69	150.31	165.85	157.07	193-193	193-191	193-193	193-191
D3S1299	0.62	151.46	166.93	158.56	202-196	202-200	196-202	202-200
D3S1279	0.85	152.3	169.6	160.19	267-267	267-275	267-267	267-275
D3S1746	0.85	153.01	169.6	/	251-263	251-255	263-251	251-255
D3S1280	0.75	153.38	170.14	161.6	215-213	215-217	213-215	215-217
D3S3531	0.73	155.49	170.14	162.79	218-220	218-216	220-216	218-216
D3S1607	0.8	158.24	172.27	164.61	234-242	234-234	242-234	234-234
D3S3575	0.79	160.63	173.34	165.61	192-190	192-190	190-190	192-190
D3S3579	0.79	161.86	173.87	166.45	188-180	188-182	180-182	188-182
D3S3668	0.83	165.72	175.47	167.47	224-224	224-224	224-224	224-224
D3S1282	0.78	170.29	180.8	170.46	168-166	168-166	166-166	168-166
D3S3523	0.68	170.93	181.87	171.02	189-193	189-197	193-197	189-197
R1	/	171.8	/	/	268-262	268-268	262-268	268-268
D3S1574	0.79	173.08	181.87	174.03	121-137	121-127	137-127	121-127

Table 4.25 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 7. Parents I.1 and I.2 have a haplotype in common (in red) which is not homozygously inherited by the two unaffected children II.1 and II.4. No DNA was available from affected individual II.3, therefore linkage could not be supported in this family. Red lines define the minimal critical interval as defined in Table 4.24. Unaff. means individual is unaffected.

success. These tissue samples were quite old and the DNA did not amplify, even after amplification with the GenomiPhi™ DNA Amplification Kit. An attempt to amplify the child's DNA was then made using *E.coli* exonuclease III, an enzyme for amplification of damaged templates like tissue extracted DNA from old paraffin sections as explained in section 2.3.2.6, but again this was not successful. Therefore, linkage to this locus could not be corroborated in this family.

4.3.1.2 Genotyping not consistent with linkage to chromosome 3q24-q26

Consanguineous Family 2 was previously found not to be consistent with linkage to the region on chromosome 3q24-q26 (see Table 4.9). Families 3, 11, 12, 13 and 15 showed no evidence of linkage to this locus. Families 4 and 16 were difficult to categorise. Linkage to this locus was excluded in Families 9 and 10 by Dr Colin Johnson and Dr Valerie Cormier-Daire, respectively.

In consanguineous Family 3, linkage was excluded because the affected child II.1 was heterozygous across the region (Table 4.27). Affected sibs II.1 and II.2 in consanguineous Family 11 were both heterozygous across the 3q24-q26 region, therefore linkage could not be supported (Table 4.28). Linkage was also excluded in distantly consanguineous Family 12 as fetus II.2 was heterozygous across the region (Table 4.29). In Family 13, parents were thought to be distantly consanguineous, therefore genotyping was performed in the parents II.1 and II.2 and affected child III.1. Linkage was excluded as no region of homozygosity was detected in affected individual III.1 (Table 4.30).

Family 3

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D3S1306	0.72	149.08	164.25	154.18	163-163	163-163	163-163
D3S1308	0.69	150.31	165.85	157.07	195-191	191-191	195-191
D3S1279	0.85	152.3	169.6	160.19	266-270	268-274	266-268
D3S1607	0.80	158.24	172.27	164.61	244-246	240-242	244-240
D3S3668	0.83	165.72	175.47	167.47	226-228	224-224	226-224
D3S1282	0.78	170.29	180.8	170.46	172-168	170-166	172-170
D3S1574	0.79	173.08	181.87	174.03	129-127	131-129	129-131

Table 4.27 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 3, showing no evidence for linkage as affected individual II.1 is heterozygous for all the markers typed (except for D3S1306, which is uninformative). Red lines define the minimal critical interval as defined in Table 4.24.

Family 11

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband	II.2 proband
D3S1306	0.72	149.08	164.25	154.18	157-163	157-157	157-157	163-157
D3S1308	0.69	150.31	165.85	157.07	191-195	191-195	191-195	191-191
D3S1746	0.85	153.01	169.6	/	267-267	251-283	267-283	267-251
D3S1607	0.80	158.24	172.27	164.61	236-244	244-234	236-234	236-244
D3S3668	0.83	165.72	175.47	167.47	222-226	236-218	222-218	222-236
D3S1282	0.78	170.29	180.8	170.46	170-168	166-166	170-166	170-166
R1	/	171.8	/	/	269-265	271-269	269-269	269-271

Table 4.28 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 11, showing no evidence for linkage. Affected individuals II.1 and II.2 are in fact both heterozygous across the region. Red lines define the minimal critical interval as defined in Table 4.24.

Family 12

Markers Chromosome 3	I.1 father	I.2 mother	II.1 unaff.	II.3 unaff.	II.2 proband
D3S1308	193-193	195-191	193-191	193-195	193-191
D3S1746	274-259	251-263	274-263	274-251	274-263
D3S1607		243-245	239-243	239-243	239-245
D3S3668	224-228	220-220	224-220	224-220	228-220
D3S1282	166-168	168-174	166-168	166-168	168-174
R1	265-275	267-269	265-267	265-267	275-269

Table 4.29 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 12, showing no evidence for linkage. Affected individual II.2 is in fact heterozygous across the region. Red lines define the minimal critical interval as defined in table 4.24. Unaff. means individual is unaffected.

Family 13

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	II.1 father	II.2 mother	III.1 proband
D3S1308	0.69	150.31	165.85	157.07	195-193	195-191	195-195
D3S1299	0.62	151.46	166.93	158.56	196-194	200-200	196-200
D3S1279	0.85	152.3	169.6	160.19	279-267	271-271	279-271
D3S1280	0.75	153.38	170.14	161.6	213-225	211-217	213-211
D3S3531	0.73	155.49	170.14	162.79	214-216	216-214	214-216
D3S1607	0.80	158.24	172.27	164.61	239-245	243-235	239-243
D3S3575	0.79	160.63	173.34	165.61	194-190	192-194	194-192
D3S3579	0.79	161.86	173.87	166.45	189-189	189-189	189-189
D3S1268	0.86	165.29	174.94	/	200-198	196-206	200-196
D3S3668	0.83	165.72	175.47	167.47	222-224	234-228	222-234
D3S1763	0.8	168.64	176.54	168.52	280-272	264-276	280-264
D3S1614	0.83	169.61	177.75	169.98	108-116	112-116	108-112
D3S1282	0.78	170.29	180.8	170.46	172-164	166-172	172-166
R1	/	171.8	/	/	262-262	266-262	262-266

Table 4.30 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 13, showing no evidence for linkage. Affected individual III.1 is heterozygous across the region, except for markers D3S3579 which is uninformative. Red lines define the minimal critical interval as defined in Table 4.24.

Linkage to this locus was also excluded in non-consanguineous Family 15. Two of the three affected fetuses (II.3 and II.4) had the same haplotype, while fetus II.2 had a different haplotype, which was the same as unaffected sib II.1 (Table 4.31).

In consanguineous Family 4, individual II.6 showed a no evidence for linkage (Table 4.32). Marker D3S1614 was heterozygous in the proband II.6. In non-consanguineous Family 16, affected child II.3 shared the same pair of haplotypes across the region as her unaffected sister II.2, but results were inconclusive as almost all the markers typed (except D3S1282) were semi-informative with at least one of the two parents being homozygous or the parents having the same haplotype (Table 4.33).

4.3.2 Chromosome 8q24 locus

Consanguineous Families 3-12 and non-consanguineous Families 13-16 were genotyped across the 8q24 region with a selection of evenly spaced markers across the region homozygous in individual VI.5 from Family 2 (see Table 4.14). Not all the markers used for typing Family 2 were used in the other families. Details of all the markers are shown in Table 2.6.

4.3.2.1 Genotyping in JATD families consistent with linkage to 8q24

Consanguineous Families 9 and 10 supported the hypothesis for linkage to 8q24 as affected children demonstrated regions of homozygosity.

In Family 9, affected individual II.1 demonstrated homozygosity between marker D8S274 and telomeric marker D8S1926 spanning about 16 cM/8 Mb (Table 4.34). DNA from the mother I.2 was of limited quantity and of poor quality, so genotyping was possible only for two markers in the region. Attempts to amplify the mother's DNA with GenomiPhi™ DNA Amplification Kit failed.

In Family 10, a region of homozygosity of about 20 cM/11 Mb was demonstrated in affected child II.10 between markers D8S256 and D8S2334 (Table 4.35). Markers

Family 15

Markers Chromosome 3	I.1 father	I.2 mother	II.1 unaff.	II.2 proband	II.3 proband	II.4 proband
D3S1308	193-193	191-193	193-191	193-191	193-193	193-193
D3S1746	251-255	259-267	251-259	251-259	255-267	255-267
D3S3668	226-226	222-230	226-222	226-222	226-230	226-230
D3S1282	168-174	170-166	168-170	168-170	174-166	174-166
R1	275-269	267-271	275-267	275-267	269-271	269-271

Table 4.31 Haplotype analysis results of chromosome 3q24-q26 region in non-consanguineous Family 15, showing no evidence for linkage. Affected fetus II.2 has a different haplotype (highlighted in pink) compared to the other two affected fetuses II.3 and II.4 (highlighted in purple). Note that fetus II.2 has the same haplotype as unaffected sib II.1. Red lines define the minimal critical interval as defined in Table 4.24. Unaff. means individual is unaffected.

Family 4

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.6 proband
D3S1764	0.79	140.46	152.62	145.53			233-233
D3S1744	0.77	148.37	161.04	153.36			242-246
D3S1306	0.72	149.08	164.25	154.18	155-159	161-161	155-161
D3S1308	0.69	150.31	165.85	157.07	193-191	193-193	193-193
D3S1279	0.85	152.3	169.6	160.19	268-268	270-266	268-270
D3S4531	0.69	152.53	169.6	160.46			257-261
D3S1607	0.8	158.24	172.27	164.61	248-234	244-234	248-244
D3S3575	0.79	160.63	173.34	165.61			180-190
D3S3668	0.83	165.72	175.47	167.47	220-224	220-220	220-220
D3S1763	0.80	168.64	176.54	168.52			227-227
D3S1614	0.83	169.61	177.75	169.98	116-112	108-110	116-108
D3S1282	0.78	170.29	180.8	170.46	166-166	166-166	166-166
D3S3523	0.68	170.93	181.87	171.02			188-190
R1	/	171.8	/	/			266-270
D3S1574	/	171.8	/	/	139-119	133-127	139-133
D3S2427	0.85	177.18	188.29	/			230-236
D3S2398	0.77	190.82	209.41	203.31			290-290

Table 4.32 Haplotype analysis results of chromosome 3q24-q26 region in consanguineous Family 4, showing no evidence for linkage. Note that markers D3S3668 and D3S1282 are uninformative. Red lines define the minimal critical interval as defined in Table 4.24.

Family 16

Markers Chromosome 3	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.2 unaff.	II.3 proband
D3S1306	0.72	149.08	164.25	154.18	163-163	157-167	163-167	163-167
D3S1308	0.69	150.31	165.85	157.07	192-192	190-190	192-190	192-190
D3S1746	0.85	153.01	169.6	/		262-270	284-270	284-270
D3S1607	0.80	158.24	172.27	164.61	240-246	236-236	240-236	240-236
D3S3668	0.83	165.72	175.47	167.47	225-235	221-221	225-221	225-221
D3S1282	0.78	170.29	180.8	170.46	168-166	172-178	168-172	168-178
D3S1574	/	171.8	/	/	129-135	129-135		

Table 4.33 Haplotype analysis results of chromosome 3q24-q26 region in non-consanguineous Family 16, showing no evidence for linkage as affected individual II.3 has the same haplotype as her unaffected sib II.2. Red lines define the minimal critical interval as defined in Table 4.24. Unaff. means individual is unaffected.

Family 9

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D8S537	0.89	135.51	150.8	147.41	163-177		163-179
D8S1100	0.65	136.77	154.02	148.81	195-183		195-192
D8S1783	0.78	136.69	154.02	148.81	172-172		172-176
D8S274	0.77	137.3	154.02	150.07	112-114		112-110
D8S272	0.8	137.41	154.02	/	247-261	247-247	247-247
D8S1837	0.81	138.91	156.59	152.44	243-231		243-243
AFM316xb1	/	138.933	/	/	190-190		190-190
AC087711	/	138.939	/	/	187-203	187-207	187-187
AC087711-2	/	138.955	/	/	250-230		250-250
AC068467	/	139.150	/	/	211-207		211-211
AC105130-R1	/	139.350	/	/	244-244		244-244
AC053480-R1	/	140.000	/	/	286-286		286-286
D8S1743	0.83	140.33	162.94	/	241-247		241-241
D8S373	/	145.74	167.9	166.08	268-260		268-268
D8S2334	/	145.73	167.9	/	156-160		156-156
D8S1926	/	145.74	167.9	166.08	252-254		252-250

Table 4.34 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 9. Patient II.1 displays homozygosity across the locus. Individual I.2 was typed only for two markers.

Family 10

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.10 proband
D8S557	0.76	132.761	145.97	142.18			204-212
D8S256	0.83	134.115	148.12	145.26			214-218
D8S1708	0.68	134.24	148.12	145.91	157-157	157-153	157-157
D8S1746	0.66	134.91	149.46	146.6	210-210	210-204	210-210
D821462	0.75	135.06	/	/	160-156	160-156	160-160
D8S1710	0.78	135.32	150.8	146.68	257-259	257-253	257-257
D8S537	0.89	135.51	150.8	147.41	163-177	163-179	163-163
D8S1100	0.65	136.77	154.02	148.81	183-189	183-189	183-183
D8S1783	0.78	136.69	154.02	148.81	172-176	172-172	172-172
D8S274	0.77	137.3	154.02	150.07	118-116	118-110	118-118
D8S272	0.8	137.41	154.02	/	239-253	239-257	239-239
D8S1837	0.81	138.91	156.59	152.44	247-243	247-233	247-247
AFM316xb1	/	138.933	/	/	190-186	190-190	190-190
AC087711	/	138.939	/	/	207-203	207-199	207-207
AC087711-2	/	138.955	/	/			252-252
AC068467	/	139.150	/	/			213-213
AC105130-R1	/	139.350	/	/			244-244
AC053480-R1	/	140.000	/	/			292-292
D8S1743	0.83	140.33	162.94	/	239-227	239-221	239-239
D8S373	/	145.74	167.9	166.08	272-264	272-276	272-272
D8S2334	/	145.73	167.9	/	157-139	155-165	157-155
D8S1926	/	145.74	167.9	166.08			246-250

Table 4.35 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 10. Patient II.10 displays homozygosity across the locus between markers D8S256 and D8S2334.

D8S1708 and D8S1746 were typed only in this family to refine the centromeric boundary of the region.

A summary of the genotyping for affected individuals in families 2, 9 and 10 is given in Table 4.36. The overlapping region of homozygosity of about 2 cM (1.65 Mb) is defined by marker D8S274 at the centromeric end and marker AC087711-2 at the telomeric end.

4.3.2.2 Genotyping not consistent with linkage to chromosome 8q24

Consanguineous Family 1 was previously found not to be consistent with linkage to this locus (see Table 4.15). Consanguineous Families 3, 4, 6, 7, 8, 11, 12, showed no evidence of linkage to 8q24. Linkage was also excluded in possibly consanguineous Family 13 as well as non-consanguineous Families 14, 15 and 16. Family 5 could not be typed at this locus because DNA from affected individual II.1 was of limited quantity.

In consanguineous Family 3, linkage was excluded because the affected child II.1 was heterozygous across the region (Table 4.38). Further typing in this family was carried out by Dr Colin Johnson and the region was definitely excluded (personal communication).

In consanguineous Family 4, no evidence for linkage was demonstrated as affected individual II.6 showed heterozygosity across the region (Table 4.39).

Markers Chromosome 8	Mb Build 33	cM Marshfield	cM deCode	Family 2 VI.3 Dutch	Family 2 VI.5 Dutch	Family 9 II.1 Italian	Family 10 II.10 Maroccan
D8S1128	128.26	139.53	135.57		238-242		
D8S1720	128.62	140.61	137.06	156-158	156-158		
D8S1732	130.04	143.29	139.58	212-214	212-214		
D8S1701	130.56	142.22	/	275-279	275-279		
AC131568	130.65	/	/	182-176	182-176		
AC103725	130.758	/	/	219-223	219-223		
AC009682	130.952	/	/	239-239	239-239		
D8S284	131.18	143.82	139.79	296-296	296-296		
D8S1712	131.323	144.36	/	225-225	225-225		
D8S557	132.761	145.97	142.18	202-202	202-202		204-212
D8S558	133.163	145.97	143.14	171-171	171-171		
D8S529	133.770	148.12	144.75	235-235	235-235		
D8S256	134.115	148.12	145.26	232-232	232-232		214-218
D8S1708	134.24	148.12	145.91				157-157
D8S1746	134.91	149.46	146.6				210-210
D8S1462	135.06	/	/	160-160	160-160		160-160
D8S1710	135.32	150.8	146.68	255-255	255-255		257-257
D8S537	135.51	150.8	147.41	180-180	180-180	163-179	163-163
D8S1100	136.77	154.02	148.81	182-182	182-182	195-192	183-183
D8S1783	136.69	154.02	148.81			172-176	172-172
D8S1761	137.25	154.02	150.07	290-290	290-290	112-110	
D8S274	137.3	154.02	150.07			112-110	118-118
D8S272	137.41	154.02	/	253-253	253-253	247-247	239-239
D8S1837	138.91	156.59	152.44	233-233	233-233	243-243	247-247
AFM316xb1	138.933	/	/	193-193	193-193	190-190	190-190
AC087711	138.939	/	/	204-204	204-204	187-187	207-207
AC087711-2	138.955	/	/	248-248	248-250	250-250	252-252
AC068467	139.15	/	/	209-209	209-211	211-211	213-213
AC105130-R1	139.35	/	/	257-257	257-245	244-244	244-244
AC053480-R1	140.00	/	/	285-285	285-287	286-286	292-292
D8S1743	140.33	162.94	/	220-220	220-246	241-241	239-239
D8S373	144.00	167.9	166.08	264-264	264-276	268-268	272-272
D8S2334	145.73	167.9	/			156-156	157-155
D8S1926	145.74	167.9	166.08	246-246	246-248	252-250	246-250

Table 4.36 Summary diagram of the genotyping in affected individuals in Families 2, 9 and 10. These data are consistent with linkage to 8q24. Homozygous typings are highlighted in bright green. Untyped markers are highlighted in white. The markers highlighted in pale blue between the two pink lines denote the overlapping region of homozygosity between the patients, flanked by markers D8S274 and AC087711-2. The ethnic origins of the families are also shown.

Family 3

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D8S558	0.86	133.163	145.97	143.14	175-175	169-175	175-169
D8S537	0.89	135.51	150.8	147.41	181-185	185-185	181-185
D8S1837	0.81	138.91	156.59	152.44	238-240	242-240	238-242

Table 4.38 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 3, showing no evidence for linkage. Affected individual II.1 is heterozygous for all the three markers typed. Pink lines define the minimal critical interval as defined in Table 4.36.

Family 4

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D8S557	**0.76**	**132.761**	**145.97**	**142.18**	207-215	205-211	207-205
D8S1710	**0.78**	**135.32**	**150.8**	**146.68**	253-259	257-253	253-257
D8S1100	**0.65**	**136.77**	**154.02**	**148.81**			182-188
D3S1837	**0.81**	**138.91**	**156.59**	**152.44**	235-243	237-233	235-237
AC087711	/	**138.939**	/	/	193-195	187-203	193-187

Table 4.39 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 4, showing no evidence for linkage. Patient II.6 is heterozygous across the region. Pink lines define the minimal critical interval as defined in Table 4.36.

In consanguineous Family 6, linkage was also excluded as no region of homozygosity was detected in affected individual II.2 (Table 4.40).

The parents I.1 and I.2 in consanguineous Family 7 did not have a haplotype in common except for allele 298 for marker D8S284 and allele 195 for marker AC087711-2, excluding the possibility of homozygosity in affected child II.3 (Table 4.41). DNA from the latter, as previously explained, was not available, therefore linkage could not be completely excluded.

The affected child in consanguineous Family 8 was heterozygous across the region (Table 4.42). Further typing in this family was carried out by Dr Colin Johnson and linkage was excluded (personal communication). Affected sibs II.1 and II.2 in consanguineous Family 11 were both heterozygous across the 8q24 region, therefore linkage was excluded (Table 4.43).

In distantly consanguineous Family 12, affected fetus II.2 showed no convincing homozygosity across the region, thus excluding linkage to this locus (Table 4.44). In Family 13 parents were thought to be distantly consanguineous, therefore genotyping was performed in the parents II.1 and II.2 and affected child III.1. Linkage was excluded as no region of homozygosity was detected in affected individual III.1 (Table 4.45). Haplotype analysis for two markers in the region was performed in non-consanguineous Family 14 and was not consistent with linkage (Table 4.46).

Linkage to this locus was also excluded in non-consanguineous Family 15. Two of the three affected fetuses (II.2 and II.4) had the same pair of haplotypes, the same as unaffected sib II.1 (Table 4.47). Fetus II.3 on the other hand, had a different pair of haplotypes compared to the other two fetuses. In non-consanguineous Family 16, affected child II.3 shared the same pair of haplotypes across the region as her unaffected sister II.2 excluding linkage to this locus (Table 4.48).

Family 6

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 proband
D8S284	0.83	131.18	143.82	139.79	296-278	278-268	296-268	296-278
D8S557	0.76	132.761	145.97	142.18	202-208	208-214	202-214	202-208
D8S256	0.83	134.115	148.12	145.26	213-213	213-213	213-213	213-213
D8S1100	0.65	136.77	154.02	148.81	191-188	188-182	191-182	191-188
D8S1837	0.81	138.91	156.59	152.44	230-230	230-230	230-230	230-230
D8S1743	0.83	140.33	162.94	/	225-245	221-221	245-221	225-221

Table 4.40 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 6. Patient II.2 displays heterozygosity across the region except for markers D8S256 and D8S1837, which are uninformative. Pink lines define the minimal critical interval as defined in Table 4.36. Unaff. means individual is unaffected.

Family 7

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.4 unaff.
D8S284	0.83	131.18	143.82	139.79	282-298	296-298	298-298	282-298
D8S557	0.76	132.761	145.97	142.18	210-216	212-212	216-212	210-212
D8S256	0.83	134.115	148.12	145.26	231-235	229-213	235-213	231-213
D8S1837	0.81	138.91	156.59	152.44	248-248	242-238	248-238	248-238
AC087711-2	/	138.955	/	/	201-195	195-187	195-187	201-187

Table 4.41 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 7. Parents I.1 and I.2 do not have alleles in common for the markers typed, with the exception of D8S284 and AC087711-2 (highlighted in red). No DNA was available from the affected individual II.3, thus linkage could not be completely substaintiated or excluded in this family. Pink lines define the minimal critical interval as defined in Table 4.36. Unaff. means individual is unaffected.

Family 8

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D8S558	0.86	133.163	145.97	143.14	167-177	175-169	167-175
D8S537	0.89	135.51	150.8	147.41	181-177	183-181	181-183
D8S1837	0.81	138.91	156.59	152.44	244-240	240-234	244-240

Table 4.42 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 8. Linkage was excluded as affected child II.1 is heterozygous across the region. Pink lines define the minimal critical interval as defined in Table 4.36.

Family 11

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband	II.2 proband
D8S284	0.83	131.18	143.82	139.79	300-300	268-298	300-268	300-268
D8S256	0.83	134.115	148.12	145.26	227-231	231-227	227-231	227-231
D8S1837	0.81	138.91	156.59	152.44	240-238	234-238	240-234	240-234
AC087711-2	/	138.955	/	/	195-195	191-195	195-191	195-191

Table 4.43 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 11, showing no evidence for linkage as affected individuals II.1 and II.2 are both heterozygous across the region. Pink lines define the minimal critical interval as defined in Table 4.36.

Family 12

Markers Chromosome 8	I.1 father	I.2 mother	II.1 unaff.	II.3 unaff.	II.2 proband
D8S529	236-254	246-248	236-248	236-248	236-246
D8S256	223-213	213-213	223-213	223-213	223-213
D8S1462	160-156	164-156	160-156	160-156	160-164
D8S1710	257-257	257-251	257-251	257-251	257-257
D8S537	180-184	180-178	180-178	180-178	180-180
D8S1108	260-268	260-264	260-264	260-264	260-260
D8S534	178-182	182-172	178-172	178-172	178-182
D8S1100	189-192	189-189	189-189	189-189	189-189
D8S1783	172-176	172-172	172-172	172-172	172-172
D8S274	110-116	108-116	110-116	110-116	110-108
D8S272		252-249	252-249	252-249	252-252
D8S1837	246-238	242-238	246-238	246-238	246-242
AC087711	197-203	203-197	197-197	197-197	197-203

Table 4.44 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 12, showing no convincing homozygosity in affected fetus II.2 across the locus, except for a small region outside the minimal critical interval from markers D8S1710 to D8S1108. Pink lines define the minimal critical interval as defined in Table 4.36. Unaff. means individual is unaffected.

Family 13

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	II.1 father	II.2 mother	III.1 proband
D8S284	0.83	131.18	143.82	139.79	278-280	270-296	278-270
D8S557	0.76	132.761	145.97	142.18	212-208	222-210	212-222
D8S558	0.86	133.163	145.97	143.14	177-163	165-167	177-165
D8S529	0.81	133.770	148.12	144.75	241-243	247-235	241-247
D8S256	0.83	134.115	148.12	145.26	227-213	213-225	227-213
D8S1462	0.75	135.06	/	/	156-164	160-156	156-160
D8S1710	0.78	135.32	150.8	146.68	258-256	258-254	258-258
D8S537	0.89	135.51	150.8	147.41	182-158	179-184	182-179
D8S1100	0.65	136.77	154.02	148.81	183-189	192-189	183-192
D8S274	0.77	137.3	154.02	150.07	112-112	118-112	112-118
D8S1837	0.81	138.91	156.59	152.44	245-239	231-243	245-231
AC087711	/	138.939	/	/	197-201	195-211	197-195

Table 4.45 Haplotype analysis results of chromosome 8q24 region in consanguineous Family 13, showing no evidence for linkage. Affected individual III.1 is heterozygous across the region. Pink lines define the minimal critical interval as defined in Table 4.36.

Family 14

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.3 proband	II.4 proband
D8S558	0.86	133.163	145.97	143.14	165-169	167-169	169-167	165-167
D8S1837	0.81	138.91	156.59	152.44	230-238	236-230	238-236	230-236

Table 4.46 Haplotype analysis results of chromosome 8q24 region in non-consanguineous Family 14, showing no evidence for linkage as affected individuals II.3 and II.4 have inherited different paternal haplotypes. Pink lines define the minimal critical interval as defined in Table 4.36.

Family 15

Markers Chromosome 8	I.1 father	I.2 mother	II.1 unaff.	II.2 proband	II.3 proband	II.4 proband
D8S529	240-258	242-252	240-242	240-242	258-252	240-242
D8S256	213-221	233-235	213-233	213-233	221-235	213-233
D8S1710	261-255	257-257	/	261-257	255-257	261-257
D8S1100	192-180	189-183	192-189	192-189	180-183	192-189
D8S1837	238-230	230-230	238-230	238-230	230-230	238-230
AC087711	197-187	195-201	197-195	197-195	187-201	197-195

Table 4.47 Haplotype analysis results of chromosome 8q24 region in non-consanguineous Family 15, showing no evidence for linkage. Affected fetus II.3 has a different haplotype (highlighted in purple) compared to the other two affected fetuses II.2 and II.4 (highlighted in pink). Note that unaffected sib II.1 has the same haplotype as fetuses II.2 and II.4. Pink lines define the minimal critical interval as defined in Table 4.36. Unaff. means individual is unaffected.

Family 16

Markers Chromosome 8	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.2 unaff.	II.3 proband
D8S1720	0.82	128.62	140.61	137.06	158-156	148-156	158-148	158-148
D8S558	0.86	133.163	145.97	143.14	167-165	163-161	167-163	167-163
D8S256	0.83	134.115	148.12	145.26	213-227	229-235	213-229	213-229
D8S1710	0.78	135.32	150.8	146.68	/	261-257	259-261	259-261
D8S1837	0.81	138.91	156.59	152.44	/	246-238	230-246	230-246
D8S1743	0.83	140.33	162.94	/	221-223	245-225	221-245	221-245

Table 4.48 Haplotype analysis results of chromosome 8q24 region in non-consanguineous Family 16, showing no evidence for linkage. Affected individual II.3 has the same haplotype as her unaffected sib II.2. Pink lines define the minimal critical interval as defined in Table 4.36. Unaff. means individual is unaffected.

4.3.3 Chromosome 12q23-q24 locus

As explained in section 4.2.5, the locus on chromosome 12q23-q24 was identified by Dr Valerie Cormier-Daire in two JATD consanguineous families. A selection of polymorphic markers were obtained from Dr Cormier-Daire and were used for genotyping consanguineous Families 1, 2, 3, 4, 6, 7, 8, 11 and 12 and non-consanguineous Families 13, 14, 15 and 16. Family 5 could not be typed at this locus because there was little DNA from affected individual II.1. Families 9 and 10 were typed by Dr Johnson and Dr Cormier-Daire, respectively. None of the above families showed evidence for linkage to 12q23-q24. Details of markers are given in Table 2.7 and their position in Table 4.49. The minimal critical interval of about 2 cM/ 3 Mb was identified between markers D12S1342 and D12S1339 (Table 4.49).

Affected individuals from consanguineous Families 1 and 2 were typed and the results were not consistent with linkage to 12q23-q24 (Table 4.50). Affected individual IV.3 in Family 1 was found to be homozygous for three consecutive markers within the minimal critical interval, but his affected cousin IV.1 was heterozygous for two of them and homozygous for a different allele for marker D12S1339. In Family 2, proband VI.3 was heterozygous for the three markers typed within the minimal critical interval (Table 4.50).

In consanguineous Family 3, affected individual II.1 was not consistent with linkage to 12q23-q24 (Table 4.51). Although she was homozygous for D12S1645, this marker was semi-informative. Moreover, she was heterozygous for a very close marker D12S1583. In consanguineous Family 4, linkage was also excluded as patient II.6 was heterozygous across the region (Table 4.52).

Affected child II.2 in Family 6 was homozygous for two consecutive markers D12S353 and D12S1605 spanning less than 2 cM (Table 4.53). The same two markers were also homozygous in the father I.1 and in the unaffected sib II.1. Although it is likely that the unaffected sib had inherited a different paternal haplotype compared to his affected sister, results in this family were inconclusive (Table 4.53).

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode
D12S346	0.84	99.46	104.65	111.02
PAH	0.8	103.17	/	/
D12S78	0.91	104.1	114.28	117.32
D12S317	0.71	105.52	114.28	117.32
D12S1342	0.83	106.99	116.08	121.28
D12S353	0.77	107.96	115.18	/
D12S1605	0.78	108.63	116.66	122.79
D12S84	0.84	108.95	117.81	123.34
D12S1583	0.87	109.71	119.55	124.65
D12S1645	0.77	109.93	119.55	/
D12S1339	0.71	109.96	118.68	/
D12S1646	0.72	114.59	121.84	/
D12S1341	0.79	114.87	123.00	/
D12S79	0.87	115.84	125.31	131.88
D12S2070	0.79	115.86	125.31	/
D12S395	0.85	119.97	136.82	140.07
D12S324	0.69	126.34	147.17	148.64

Table 4.49 Markers on chromosome 12q23-q24 used for haplotype analysis. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval identified in two consanguineous JATD families by Dr Cormier-Daire.

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5
D12S346	0.84	99.46	104.65	111.02	187-197	195-197	189-195	191-191
PAH	0.8	103.17	/	/	241-253	241-245	244-244	236-256
D12S353	0.77	107.96	115.18	/	178-186	174-188	178-188	176-176
D12S84	0.84	108.95	117.81	123.34	214-224	200-200	202-212	210-216
D12S1583	0.87	109.71	119.55	124.65	219-241	239-239	237-241	235-235
D12S1339	0.71	109.96	118.68	/	266-266	268-268	269-271	269-271
D12S79	0.87	115.84	125.31	131.88	162-162	156-162	154-156	154-156
D12S2070	0.79	115.86	125.31	/	98-98	88-98	84-88	88-88
D12S395	0.85	119.97	136.82	140.07	244-248	240-248	240-240	224-232
D12S324	0.69	126.34	147.17	148.64	242-246	242-246	242-244	238-244

Table 4.50 Haplotype analysis results on chromosome 12q23-q24 in patients from consanguineous Families 1 and 2. In Family 1, affected child IV.3 is homozygous for two consecutive markers, D12S84 and D12S1583 within the minimal critical interval, although his affected cousin (IV.1) is heterozygous for the same two markers. Both affected cousins, IV.1 and IV.3, are homozygous for D12S1339, but for different alleles. In Family 2, proband VI.5 is homozygous for two non-consecutive markers, D12S353 and D12S1583, and his affected sib VI.3 is heterozygous for both of them. Therefore linkage was excluded in both families. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval.

Family 3

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D12S346	0.84	99.46	104.65	111.02	202-190	200-184	202-200
D12S353	0.77	107.96	115.18	/	186-178	178-178	186-178
D12S84	0.84	108.95	117.81	123.34	216-218	198-212	216-198
D12S1583	0.87	109.71	119.55	124.65	237-243	239-239	237-239
D12S1645	0.77	109.93	119.55	/	213-213	213-217	213-213
D12S1339	0.71	109.96	118.68	/	269-273	269-271	269-269
D12S1646	0.72	114.59	121.84	/	257-257	249-251	257-249
D12S1341	0.79	114.87	123.00	/	204-208	210-216	204-210
D12S79	0.87	115.84	125.31	131.88	160-154	160-158	160-160

Table 4.51 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 3. Linkage in this family could not be supported. Affected individual II.1 is homozygous for D12S1645 although this marker is semi-informative and she is heterozygous for a very close marker, D12S1583. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval.

Family 4

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.6 proband
D12S346	0.84	99.46	104.65	111.02	192-188	194-190	192-194
PAH	0.8	103.17	/	/	236-240	240-244	236-240
D12S78	0.91	104.1	114.28	117.32	197-191	189-181	197-189
D12S317	0.71	105.52	114.28	117.32	250-270	250-270	250-250
D12S1342	0.83	106.99	116.08	121.28	277-273	281-271	277-281
D12S353	0.77	107.96	115.18	/	184-182	184-186	184-184
D12S1605	0.78	108.63	116.66	122.79	196-202	204-198	196-204
D12S84	0.84	108.95	117.81	123.34	215-209	217-219	215-217
D12S1583	0.87	109.71	119.55	124.65	242-221	221-234	242-221
D12S1339	0.71	109.96	118.68	/	268-270	266-266	268-266
D12S79	0.87	115.84	125.31	131.88	174-156	162-168	174-162

Table 4.52 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 4. Linkage in this family could not be supported. Affected individual II.6 is homozygous for D12S353 within the region, although he is heterozygous for the two flanking markers, D12S1342 and D12S1605. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval.

Family 6

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 proband
D12S78	0.91	104.1	114.28	117.32	193-185	173-183	185-173	193-173
D12S317	0.71	105.52	114.28	117.32	270-270	250-250	270-250	270-250
D12S1342	0.83	106.99	116.08	121.28	281-279	273-281	279-273	281-273
D12S353	0.77	107.96	115.18	/	176-176	176-188	176-176	176-176
D12S1605	0.78	108.63	116.66	122.79	200-200	200-202	200-200	200-200
D12S84	0.84	108.95	117.81	123.34	213-215	219-221	215-219	213-219
D12S1583	0.87	109.71	119.55	124.65	225-225	223-225	225-223	225-223
D12S1339	0.71	109.96	118.68	/	265-267	267-269	267-267	265-267

Table 4.53 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 6. Patient II.2 displays a short stretch of homozygosity within the minimal critical interval but both markers are semi-informative and unaffected sib II.1 is also homozygous for the same two markers. Therefore, linkage in this family could not be supported. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval. Unaff. means individual is unaffected.

Family 7

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.4 unaff.
D12S353	0.77	107.96	115.18	/	172-178	176-178	172-176	178-178
D12S84	0.84	108.95	117.81	123.34	219-213	217-213	219-217	213-213
D12S1583	0.87	109.71	119.55	124.65	225-225	235-241	225-235	225-241
D12S1339	0.71	109.96	118.68	/	265-267	275-267	265-275	267-267

Table 4.54 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 7 are not consistent with linkage. Parents I.1 and I.2 have alleles in common for markers D12S353, D12S84 and D12S1339 (shown in red), although unaffected individual II.4 is homozygous for those markers. No DNA was available from affected individual II.3. Herozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval. Unaff. means individual is unaffected.

In consanguineous Family 7, haplotype analysis was not consistent with linkage as parents I.1 and I.2 do not have an allele in common at D12S1583 and unaffected child II.4 is homozygous for D12S353, D12S84 and D12S1339 (Table 4.54). No DNA from the affected, deceased child (II.3) was available for typing.

Affected individual II.1 in consanguineous Family 8 was homozygous only for D12S84 within the minimal critical interval and heterozygous for the two flanking markers D12S353 and D12S1583 (Table 4.55). Linkage could not be supported in this family. In consanguineous Family 11, no evidence for linkage could be corroborated as the two affected children II.1 and II.2 were heterozygous across the interval (Table 4.56).

In distantly consanguineous Family 12, affected fetus II.2 showed no evidence for linkage to the region (Table 4.57). Fetus II.2 had the same pair of haplotypes as his unaffected sister II.1 across the region and was homozygous for D12S353, but so was her sister.

No evidence for linkage could be demonstrated in distantly consanguineous Family 13 as no region of homozygosity was detected in affected individual III.1 (Table 4.58). In non-consanguineous Family 14, affected children II.3 and II.4 did not have the same haplotypes across the region, therefore linkage could not be supported (Table 4.59). The three affected fetuses II.2, II.3 and II.4 in non-consanguineous Family 15 did not have the same haplotypes in the region, therefore linkage could not be demonstrated (Table 4.60). Haplotype analysis results in non-consanguineous Family 16 were inconclusive (Table 4.61). Although the affected child II.3 had the same haplotype as her unaffected sister, possibly excluding linkage, most of the markers were uninformative in the father.

Family 8

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D12S346	0.84	99.46	104.65	111.02	190-190	188-198	190-188
D12S353	0.77	107.96	115.18	/	178-184	176-176	178-176
D12S84	0.84	108.95	117.81	123.34	200-210	200-216	200-200
D12S1583	0.87	109.71	119.55	124.65	/	220-239	239-220
D12S1339	0.71	109.96	118.68	/	267-269	265-273	267-265
D12S79	0.87	115.84	125.31	131.88	162-172	156-156	162-156

Table 4.55 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 8. Linkage in this family could not be supported, although affected individual II.1 is homozygous for D12S84 within the interval, and could be homozygous for a small region (2 cM) between markers D12S353 and D12S1583. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval.

Family 11

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband	II.2 proband
D12S353	0.77	107.96	115.18	/	186-176	172-178	186-172	186-172
D12S84	0.84	108.95	117.81	123.34	199-215	213-201	199-213	199-213
D12S1583	0.87	109.71	119.55	124.65	243-237	241-241	243-241	243-241
D12S1339	0.71	109.96	118.68	/	267-267	275-267	267-275	267-275

Table 4.56 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 11 are not consistent with linkage. Both affected individuals II.1 and II.2 are heterozygous across the region. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval.

Family 12

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.3 unaff.	II.2 proband
D12S353	0.77	107.96	115.18	/	177-185	177-179	177-177	185-177	177-177
D12S84	0.84	108.95	117.81	123.34	215-199	219-217	215-219	199-219	215-219
D12S1583	0.87	109.71	119.55	124.65	240-224	236-242	240-236	224-236	240-236
D12S1339	0.71	109.96	118.68	/	268-276	270-268	268-270	276-270	268-270

Table 4.57 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 12 are not consistent with linkage to this locus. D12S353 is homozygous in patient II.2 as well as in unaffected sib II.1. Heterozygosity, physical and genetic distances of each marker are indicated. Blue lines define the minimal critical interval. Unaff. means individual is unaffected.

Family 13

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	II.1 father	II.2 mother	III.1 proband
D12S346	0.84	99.46	104.65	111.02	188-196	198-180	188-198
D12S353	0.77	107.96	115.18	/	188-176	186-178	188-186
D12S84	0.84	108.95	117.81	123.34	212-216	210-212	212-210
D12S1583	0.87	109.71	119.55	124.65	223-235	237-235	223-237
D12S1339	0.71	109.96	118.68	/	265-267	267-267	265-267
D12S79	0.87	115.84	125.31	131.88	154-156	168-154	154-168

Table 4.58 Haplotype analysis results on chromosome 12q23-q24 in consanguineous Family 13 showing no evidence for linkage as affected individual III.1 is heterozygous across the region. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval.

Family 14

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.3 proband	II.4 proband
D12S353	0.77	107.96	115.18	/	178-178	186-184	178-186	178-184
D12S84	0.84	108.95	117.81	123.34	223-213	213-213	223-213	213-213
D12S1583	0.87	109.71	119.55	124.65	221-225	225-241	221-225	225-225
D12S1339	0.71	109.96	118.68	/	266-270	266-280	266-266	270-266

Table 4.59 Haplotype analysis results on chromosome 12q23-q24 in non-consanguineous Family 14 are not consistent with linkage. Affected individuals II.3 and II.4 have inherited different haplotypes from the parents. Heterozygosity, physical and genetic positions for each marker are given. Blue lines define the minimal critical interval.

Family 15

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 proband	II.3 proband	II.4 proband
D12S353	0.77	107.96	115.18	/	179-177	179-177	179-177	177-177	177-177	179-179
D12S84	0.84	108.95	117.81	123.34	217-215	201-213	217-213	215-213	215-213	217-201
D12S1583	0.87	109.71	119.55	124.65	238-240	221-238	238-238	240-238	240-221	238-221
D12S1645	0.77	109.93	119.55	/	213-213	219-213	213-213	213-213	213-219	213-219
D12S1339	0.71	109.96	118.68	/	268-268	266-266	268-266	268-266	268-266	268-266

Table 4.60 Haplotype analysis results on chromosome 12q23-q24 in non-consanguineous Family 15 are not consistent with linkage to this locus. The three affected fetuses II.2, II.3 and II.4 have different haplotypes in the region. Heterozygosity, physical and genetic distances of each marker are indicated. Blue lines define the minimal critical interval. Unaff. means individual is unaffected.

Family 16

Markers Chromosome 12	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.2 unaff.	II.3 proband
D12S346	0.84	99.46	104.65	111.02	189-201	181-199	189-181	189-199
D12S353	0.77	107.96	115.18	/	175-175	173-181	175-173	175-173
D12S1605	0.78	108.63	116.66	122.79	200-204	204-200	200-204	200-204
D12S84	0.84	108.95	117.81	123.34	215-215	199-215	215-199	215-199
D12S1583	0.87	109.71	119.55	124.65	238-238	240-236	238-240	238-240
D12S1645	0.77	109.93	119.55	/	212-212	212-216	212-212	212-212
D12S1339	0.71	109.96	118.68	/	266-266	266-266	266-266	266-266

Table 4.61 Haplotype analysis results on chromosome 12q23-q24 in non-consanguineous Family 16. Affected individual II.3 and her unaffected sib II.2 have the same haplotypes except for marker D12S346 at which they have inherited different maternal alleles. Note how the father I.1 is uninformative for most of the markers. Heterozygosity, physical and genetic distances of each marker are indicated. Blue lines define the minimal critical interval. Unaff. means individual is unaffected.

4.3.4 Chromosome 15q13 locus

As explained in section 4.2.5, the locus on chromosome 15q13 was identified by Dr Colin Johnson and supported by data from five consanguineous families and three non-consanguineous families (Morgan *et al.*, 2003). Families A, C and D, reported in the paper by Morgan and collaborators, correspond to Families 3, 8 and 9, respectively, as described in this study (Morgan *et al.*, 2003).

Subsequent to the publication of the paper, more typing was carried out in Families 3 and 8 with the aim of refining further the boundaries of the region. Consanguineous Families 3-12 and non-consanguineous Families 13-16 were also genotyped across the 15q13 locus with a selection of evenly spaced markers across the region. A few polymorphic markers were obtained from Dr Colin Johnson and others were designed as explained in section 2.3.2.2. Details of all the markers are given in Table 2.8.

4.3.4.1 Families 3 and 8

Results described by Morgan and collaborators have been summarised in Table 4.62. The minimal critical interval was 1.2 cM (~1.5 Mb) in size between markers D15S165 and D15S1010 (Morgan *et al.*, 2003). In this study, PCRs to amplify almost all markers typed by Morgan and collaborators on 15q13 were repeated and run again in Families 3 and 8. Further finely spaced markers were also typed in these two families. The results of haplotype analysis in Family 3 are shown in Table 4.63. A large homozygous region (22 cM/7 Mb) was demonstrated in affected child II.1 from D15S817 to D15S165, although markers D15S976 and D15S1031 gave different results compared to those published. The child was in fact heterozygous for both markers (Table 4.63). Therefore, assuming linkage, the region was narrowed down in this family at the telomeric end to D15S165.

In Family 8, the affected child II.1 showed homozygosity from D15S1019 to GATA50C03, spanning about 24 cM/11 Mb, although only two markers, D15S165 and D15S1031, were fully informative (Table 4.64). These results were then collated with

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	Proband A Fam 3 (II.1)	Proband B	Proband C Fam 8 (II.1)	Proband D Fam 9 (II.1)	Proband E
D15S128	0.78	22.54	6.11	5.82	202-202	192-192	198-198	198-206	198-198
D15S822	0.79	24.82	12.3	/	260-260	260-260	260-284	286-306	264-264
D15S1048	0.84	27.44	19.12	21.07	223-223	229-229	227-227	211-231	225-225
D15S165	0.77	28.84	20.24	22.64	199-199	199-199	193-193	197-201	197-197
D15S976	0.63	29.13	21.58	22.65	150-150	150-150	144-144	150-150	154-154
D15S1013	0.53	29.31	21.58	24.63	168-168	178-178	168-168	168-168	168-168
D15S1031	0.74	29.66	21.58	24.63	229-229	303-303	303-303	305-305	303-303
D15S1010	0.8	30.66	23.89	/	227-233	229-229	227-227	213-213	213-213
D15S231	0.5	30.82	24.06	26.48	108-112	108-108	108-108	108-108	108-108
D15S1007	0.87	31.31	25.86	28.86	174-176	176-186	184-184	182-176	174-188
GATA50C03	0.72	/	34.83	/	272-272	272-272	272-272	280-280	276-284
D15S659	0.84	43.95	43.47	44.5	197-173	197-197	177-201	186-186	178-198
D15S643		57.29	53.33	59.51	205-213	221-221	221-221	203-209	213-213
D15S1507					208-204	204-204	212-216	204-204	208-208

Table 4.62 Summary of haplotype results published by Morgan and colleagues (Morgan *et al*., 2003) in five consanguineous JATD families. Only the haplotypes of the affected children are shown. Green lines define the minimal critical interval between D15S165 and D15S1010.

Family 3

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D15S817	0.79	22.02	4.78	/			164-164
D15S1021	0.79	22.43	4.78	4.84	222-222	222-224	222-222
D15S128	0.78	22.54	6.11	5.82			202-202
D15S822	0.79	24.82	12.30	/			260-260
D15S1002	0.79	25.32	14.58	15.05	148-135	148-135	148-148
D15S219	/	25.37	14.58	15.05	223-223	223-219	223-223
D15S156	0.51	25.38	14.58	15.05	238-238	238-250	238-238
D15S217	/	25.54	/	/			246-246
AC079090	/	25.66	/	/	242-238	242-246	242-242
AC024474	/	26.68	/	/	215-211	215-205	215-215
D15S1019	0.63	27.24	19.12	20.59	220-220	220-222	220-220
AC090763	/	27.28	/	/	262-260	262-260	262-262
D15S1048	0.84	27.44	19.12	21.07			223-223
AC022613	/	27.50	/	/	273-275	273-273	273-273
D15S1043	0.52	27.81	20.24	21.57	231-231	231-231	231-231
D15S165	0.77	28.84	20.24	22.64			199-199
D15S976	0.63	29.13	21.58	22.65	148-148	150-148	148-150
D15S1013	0.53	29.31	21.58	24.63			168-168
D15S1031	0.74	29.66	21.58	24.63	245-261	247-245	245-247
D15S1010	0.80	30.66	23.89	/			227-233
D15S231	0.50	30.82	24.06	26.48			108-112
D15S1007	0.87	31.31	25.86	28.86			167-171
ACTC	0.87	32.66	31.46	/			69-89
GATA50C03	0.72	/	/	/			284-272
D15S659	0.84	43.95	43.47	44.5			197-173

Table 4.63 Haplotype analysis results on chromosome 15q13 in consanguineous Family 3. The region of homozygosity in patient II.1 does not extend as telomeric as D15S1010 as originally reported by Morgan and colleagues (Morgan *et al*., 2003). The results which were found to differ from those published in proband II.1 are highlighted in yellow. Not all the markers were typed in the unaffected parents. Heterozygosity, physical and genetic distances of each marker are indicated. Markers included in the paper by Morgan and colleagues are shown in bold.

Family 8

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband
D15S817	0.79	22.02	4.78	/	158-154	158-150	158-158
D15S128	0.78	22.54	6.11	5.82	198-202	198-198	198-198
D15S822	0.79	24.82	12.30	/	260-292	284-264	260-284
D15S1002	0.79	25.32	14.58	15.05	135-148	135-148	135-135
D15S219	/	25.37	14.58	15.05	203-203	223-227	203-223
D15S156	0.51	25.38	14.58	15.05	238-259	240-256	238-240
D15S217	/	25.54	/	/	237-237	245-237	237-245
AC079090	/	25.66	/	/	238-240	228-238	238-228
AC024474	/	26.68	/	/	209-213	211-205	209-211
D15S1019	0.63	27.24	19.12	20.59	220-220	220-230	220-220
AC090763	/	27.28	/	/	262-262	262-262	262-262
D15S1048	0.84	27.44	19.12	21.07	226-220	226-226	226-226
AC022613	/	27.50	/	/	273-273	273-273	273-273
D15S1043	0.52	27.81	20.24	21.57	231-231	231-231	231-231
D15S165	0.77	28.84	20.24	22.64	193-197	193-195	193-193
D15S976	0.63	29.13	21.58	22.65	142-148	142-142	142-142
D15S1031	0.74	29.66	21.58	24.63	249-253	249-245	249-249
D15S1010	0.80	30.66	23.89	/	227-227	227-227	227-227
D15S231	0.50	30.82	24.06	26.48	108-108	108-112	108-108
AC019278	/	31.00	/	/	192-192	192-203	192-192
D15S995	0.67	31.24	25.86	28.86	242-242	242-244	242-242
D15S1007	0.87	31.31	25.86	28.86	177-177	177-177	177-177
AC087638	/	31.59	/	/	280-280	280-316	280-280
D15S1232	/	32.56	31.46	/	319-319	319-315	319-319
ACTC	0.87	32.66	31.46	/	91-91	91-71	91-91
D15S971	0.82	32.96	31.46	32.71	306-306	306-300	306-306
D15S1042	0.82	33.84	32.58	/	286-286	286-274	286-286
GATA50C03	0.72	/	/	/	280-280	280-276	280-280
D15S994	0.72	38.16	40.25	/	164-171	171-166	164-171
D15S659	0.84	43.95	43.47	44.5	201-176	176-197	201-176

Table 4.64 Haplotype analysis results on chromosome 15q13 in consanguineous Family 8. The region of homozygosity in proband II.1 from D15S1019 to GATA50C03 is about 24 cM/11 Mb in size. Note how almost all the markers homozygous in the proband are semi-informative except for D15S165 and D15S1031. Heterozygosity, physical and genetic distances of each marker are indicated. Markers included in the paper by Morgan and colleagues are shown in bold (Morgan *et al*., 2003).

the previous ones reported by Morgan and collaborators and it was noted that the homozygous regions in the affected children from Families 3 and 9 no longer overlapped (Table 4.65).

Following genotyping in Families 3 and 8 carried out in this study and assuming linkage, the candidate region of 2 cM/2.5 Mb was refined between markers AC024474 at the centromeric end and D15S976 at the telomeric end (Table 4.65). However, patient II.1 in Family 8 was also homozygous for a 19 cM/25 Mb region on chromosome 3q24-q26 (see Table 4.22), while data from the affected child II.1 in Family 9 was also consistent with linkage to the 8q24 locus, showing a 16 cM/8 Mb region of homozygosity (see Table 4.34). The stretch of homozygosity in proband II.1 in Family 9 on 15q13 was about 6 cM/2.5 Mb and did not overlap with the homozygous region in Family 3.

4.3.4.2 Genotyping not consistent with linkage to chromosome 15q13

Genotyping across the 15q13 locus was performed in consanguineous Families 1, 2, 4, 5, 6, 7, 11 and 12 and non-consanguineous Families 13, 14, 15 and 16. Family 10 was typed by Dr Valerie Cormier-Daire. None of the above families were consistent with linkage to this locus. Results of haplotype analysis obtained in patients from Families 1, 2, 4 and 5 are shown in Table 4.66 and were not consistent with linkage. Affected individuals from Families 1, 2, 3, 4 and 8 all shared identical homozygous alleles at marker D15S1013, although the heterozygosity of this marker was 0.53. Therefore D15S1013 was not used for typing in further families.

In consanguineous Family 6, the proband II.2 was homozygous for four markers within the minimal critical interval, although only two, AC090763 and D15S165, were fully informative (Table 4.67). Furthermore, the small potential region of homozygosity in the affected child II.2 was interrupted by heterozygous marker AC022613. Therefore the results in II.2 were not consistent with linkage to 15q13.

Consanguineous Family 7 could be consistent with linkage to this locus, but no DNA from the affected deceased child (II.3) was available to corroborate the hypothesis of

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	Proband A Fam 3 (II.1)	Proband B	Proband C Fam 8 (II.1)	Proband D Fam 9 (II.1)	Proband E
D15S817	0.79	22.02	4.78	/	164-164		158-158		
D15S1021	0.79	22.43	4.78	4.84	222-222				
D15S128	0.78	22.54	6.11	5.82	202-202	192-192	198-198	198-206	198-198
D15S822	0.79	24.82	12.30	/	260-260	260-260	260-284	286-306	264-264
D15S1002	0.79	25.32	14.58	15.05	148-148		135-135		
D15S219	/	25.37	14.58	15.05	223-223		203-223		
D15S156	0.51	25.38	14.58	15.05	238-238		238-240		
D15S217	/	25.54	/	/	246-246		237-245		
AC079090	/	25.66	/	/	242-242		238-228		
AC024474	/	26.68	/	/	215-215		209-211		
D15S1019	0.63	27.24	19.12	20.59	220-220		220-220		
AC090763	/	27.28	/	/	262-262		262-262		
D15S1048	0.84	27.44	19.12	21.07	223-223	229-229	226-226	211-231	225-225
AC022613	/	27.50	/	/	273-273		273-273		
D15S1043	0.52	27.81	20.24	21.57	231-231		231-231		
D15S165	0.77	28.84	20.24	22.64	199-199	199-199	193-193	197-201	197-197
D15S976	0.63	29.13	21.58	22.65	148-150	150-150	142-142	150-150	154-154
D15S1013	0.53	29.31	21.58	24.63	168-168	178-178	168-168	168-168	168-168
D15S1031	0.74	29.66	21.58	24.63	245-247	303-303	249-249	305-305	303-303
D15S1010	0.80	30.66	23.89	/	227-233	229-229	227-227	213-213	213-213
D15S231	0.50	30.82	24.06	26.48	108-112	108-108	108-108	108-108	108-108
AC019278	/	31.00	/	/			192-192		
D15S995	0.67	31.24	25.86	28.86			242-242		
D15S1007	0.87	31.31	25.86	28.86	167-171	176-186	177-177	182-176	174-188
AC087638	/	31.59	/	/			280-280		
D15S1232	/	32.56	31.46	/			319-319		
ACTC	0.87	32.66	31.46	/	69-89		91-91		
D15S971	0.82	32.96	31.46	32.71			306-306		
D15S1042	0.82	33.84	32.58	/			286-286		
GATA50C03	0.72	/	/	/	284-272	272-272	280-280	280-280	276-284
D15S994	0.72	38.16	40.25	/			164-171		
D15S659	0.84	43.95	43.47	44.5	197-173	197-197	201-176	186-186	178-198

Table 4.65 Summary of haplotype analysis results on chromosome 15q13 after further typing in consanguineous Families 3 and 8. Only the haplotypes from the affected children are shown. Note the heterozygous results for markers D15S976 and D15S1031 (highlighted in yellow) in patient II.1 in Family 3, differ from the previously reported data, thus the homozygous region in Family 3 does not overlap with the one in Family 9. Orange lines define the minimal critical interval determined by Families 3 and 8. Heterozygosity, physical and genetic distances of each marker are indicated. Markers included in the paper by Morgan and colleagues are shown in bold (Morgan *et al.*, 2003).

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5	Family 4 II.6	Family 5 II.1
D15S817	0.79	22.02	4.78	/	154-154	154-154	150-154		158-158	164-174
D15S128	0.78	22.54	6.11	5.82	200-200	202-206	198-200	200-204	202-204	202-204
D15S822	0.79	24.82	12.3	/	259-287	259-294	263-287	263-291	259-267	279-295
D15S217	/	25.54	/	/	234-250	218-250	238-246	246-250	230-246	242-246
D15S1048	0.84	27.44	19.12	21.07	220-226	224-226	218-220	220-228	200-208	218-220
D15S165	0.77	28.84	20.24	22.64	183-199	183-205	183-199	183-183	197-201	183-201
D15S976	0.63	29.13	21.58	22.65	142-142	142-148	142-148	142-142	148-150	148-148
D15S1013	0.53	29.31	21.58	24.63	168-168	168-168	168-168	168-168	168-168	176-176
D15S1031	0.74	29.66	21.58	24.63	246-246	246-250	261-261	249-261	249-253	249-249
D15S1010	0.8	30.66	23.89	/	229-233	225-229	229-233	227-231	225-229	229-233
D15S231	0.5	30.82	24.06	26.48	107-107	107-107	107-111	107-111	103-107	107-111
D15S1007	0.87	31.31	25.86	28.86	169-175	169-177	163-169	169-179	163-169	169-177
ACTC	0.87	32.66	31.46	/	83-87	83-89	69-89	80-89	83-83	69-91
GATA50C03	0.72	/	/	/	280-280	276-280	280-284	276-284	276-292	276-284
D15S659	0.84	43.95	43.47	44.5	176-176	176-196	180-200	180-184		

Table 4.66 Haplotype analysis results on chromosome 15q13 in affected individuals in consanguineous Families 1, 2, 4 and 5. None of the children show homozygosity across the region determined by the results shown in Table 4.65. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the new minimal critical interval determined by the results shown in Table 4.65.

Family 6

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 proband
D15S128	0.78	22.54	6.11	5.82	198-204	204-198	204-198	198-204
D15S822	0.79	24.82	12.30	/	283-287	287-279	287-279	283-287
D15S1002	0.79	25.32	14.58	15.05	135-149	135-137	135-137	135-135
D15S156	0.51	25.38	14.58	15.05	260-256	238-238	260-238	260-238
D15S217	/	25.54	/	/	217-237	241-245	217-245	217-241
AC079090	/	25.66	/	/	240-234	240-228	240-228	240-240
AC024474	/	26.68	/	/	217-209	211-209	217-209	217-211
D15S1019	0.63	27.24	19.12	20.59	220-220	220-220	220-220	220-220
AC090763	/	27.28	/	/	263-269	263-261	263-261	263-263
AC022613	/	27.50	/	/	272-272	274-274	272-274	272-274
D15S1043	0.52	27.81	20.24	21.57	231-223	231-231	231-231	231-231
D15S165	0.77	28.84	20.24	22.64	197-183	197-183	197-183	197-197
D15S1031	0.74	29.66	21.58	24.63	249-249	249-261	249-261	249-249
D15S1010	0.80	30.66	23.89	/	229-227	227-233	229-233	229-227
D15S1007	0.87	31.31	25.86	28.86	161-165	165-175	161-175	161-165

Table 4.67 Haplotype analysis results on chromosome 15q13 in consanguineous Family 6. Proband II.2 was not consistent with linkage to this locus as she was heterozygous for marker AC022613. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65. Unaff. means individual is unaffected.

linkage. Haplotype analysis was performed in the parents (I.1 and I.2) and unaffected sibs (II.1 and II.4) and was consistent with linkage (Table 4.68). The parents in fact had a haplotype in common across the region, thus the affected child could in principle be homozygous. Only one marker in the minimal critical interval, AC090763, was fully informative and the two unaffected children were heterozygous for it.

Consanguineous Family 11 was not consistent with linkage to 15q13 as both affected children II.1 and II.2 were heterozygous for two markers in the region (Table 4.69). In Family 12, linkage could not be supported, as the affected child II.2 was homozygous only for D15S165 within the critical interval (Table 4.70). In Family 13, assuming the parents were distantly related, linkage could not be supported, as patient III.3 was heterozygous across the region (Table 4.71).

In non-consanguineous Family 14 linkage could not be supported, as the two affected sibs did not share the same pair of haplotypes across the region (Table 4.72). Linkage to this locus was also excluded in non-consanguineous Family 15. Two of the three affected fetuses (II.2 and II.4) had the same haplotypes while the third affected fetus II.3 had a different pair of haplotypes, the same as the unaffected sib II.1 (Table 4.73). Non-consanguineous Family 16 was also not consistent with linkage to 15q13, as the affected child II.3 had the same pair of haplotypes as her unaffected sister (Table 4.74).

4.3.5 Chromosome 14q24-q32 locus

The locus on chromosome 14q24-q32 was identified in consanguineous Family 4 following a genome-wide screen in patient II.6. The homozygous region in probands II.6 and II.8 was about 19 cM/14 Mb in size between markers D14S74 and D14S977 (see Table 4.19). Haplotype analysis was performed in non-consanguineous Family 15 across this locus. Results are shown in Table 4.75 and were not consistent with linkage.

Family 7

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.4 unaff.
D15S1021	0.79	22.43	4.78	4.84	221-221	221-233	221-221	221-221
D15S128	0.78	22.54	6.11	5.82	198-206	206-200	198-206	198-206
D15S1513	0.61	23.60	12.3	8.60	227-231	231-231	227-231	227-231
D15S822	0.79	24.82	12.30	/	260-286	286-260	260-286	260-286
D15S1002	0.79	25.32	14.58	15.05	147-135	135-135	147-135	135-135
D15S219	/	25.37	14.58	15.05	223-227	227-203	223-227	227-203
D15S156	0.51	25.38	14.58	15.05	238-238	238-238	238-238	238-238
D15S217	/	25.54	/	/	217-237	237-245	217-237	237-245
AC079090	/	25.66	/	/	238-238	238-228	238-238	238-228
AC024474	/	26.68	/	/	207-207	207-209	207-207	207-209
D15S1019	0.63	27.24	19.12	20.59	222-220	220-220	220-222	220-220
AC090763	/	27.28	/	/	261-257	257-259	261-257	257-259
D15S1048	0.84	27.44	19.12	21.07	222-220	220-220	222-220	220-220
AC022613	/	27.50	/	/	274-274	274-274	274-274	274-274
D15S1043	0.52	27.81	20.24	21.57	231-231	231-231	231-231	231-231
D15S1031	0.74	29.66	21.58	24.63	245-253	253-245	245-253	253-245
D15S1010	0.80	30.66	23.89	/	229-233	233-227	229-233	233-227
D15S1007	0.87	31.31	25.86	28.86	169-171	171-183	169-171	171-183

Table 4.68 Haplotype analysis results on chromosome 15q13 in consanguineous Family 7. The parents I.1 and I.2 have a haplotype in common (in red) which could have been homozygously inherited by the affected child II.3, but no DNA was available from him. Therefore linkage could not be supported in this family. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65. Unaff. means individual is unaffected.

Family 11

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 proband	II.2 proband
D15S1021	0.79	22.43	4.78	4.84	233-233	221-233	233-221	233-233
D15S128	0.78	22.54	6.11	5.82	198-198	198-200	198-200	198-200
D15S1513	0.61	23.60	12.3	8.60	231-231	231-235	231-235	231-235
D15S822	0.79	24.82	12.30	/	256-260	282-290	256-290	256-290
D15S1002	0.79	25.32	14.58	15.05	135-149	135-135	135-135	135-135
D15S219	/	25.37	14.58	15.05	223-223	223-221	223-221	223-221
D15S156	0.51	25.38	14.58	15.05	238-238	238-238	238-238	238-238
D15S217	/	25.54	/	/	245-229	217-245	245-245	245-245
AC079090	/	25.66	/	/	238-234	228-228	238-228	238-228
AC024474	/	26.68	/	/	207-207	207-215	207-215	207-215
D15S1019	0.63	27.24	19.12	20.59	220-220	220-220	220-220	220-220
AC090763	/	27.28	/	/	270-266	262-264	270-264	270-264
D15S1048	0.84	27.44	19.12	21.07	222-222	220-222	222-222	222-222
AC022613	/	27.50	/	/	272-274	272-272	272-272	272-272
D15S1043	0.52	27.81	20.24	21.57	221-231	231-231	221-231	221-231
D15S1031	0.74	29.66	21.58	24.63	249-249	249-249	249-249	249-249
D15S1010	0.80	30.66	23.89	/	233-229	211-227	233-227	233-227

Table 4.69 Haplotype analysis results on chromosome 15q13 in consanguineous Family 11. Patients II.1 and II.2 are heterozygous for two markers within the critical region, AC090763 and D15S1043. Therefore linkage could not be supported in this family. Note how almost all the markers are semi- or uninformative. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65.

Family 12

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.3 unaff.	II.2 proband
D15S128	**0.78**	**22.54**	**6.11**	**5.82**	205-207	201-209	205-209	205-201	205-201
D15S822	**0.79**	**24.82**	**12.30**	/	290-282	282-286	290-282	290-282	290-286
D15S219	/	**25.37**	**14.58**	**15.05**	233-203	225-223	233-225	233-225	233-225
AC079090	/	**25.66**	/	/	228-228	228-238	228-228	228-228	228-238
AC024474	/	**26.68**	/	/	211-215	211-211	211-211	211-211	211-211
D15S1048	**0.84**	**27.44**	**19.12**	**21.07**	222-226	222-226	222-222	222-222	222-226
D15S165	**0.77**	**28.84**	**20.24**	**22.64**	183-197	193-183	183-193	183-193	183-183
D15S1031	**0.74**	**29.66**	**21.58**	**24.63**	307-303	303-299	307-303	307-303	307-299

Table 4.70 Haplotype analysis results on chromosome 15q13 in consanguineous Family 12 are not consistent with linkage to this locus. Fetus II.2 is homozygous for D15S165, although heterozygous for marker D15S1048. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65. Unaff. means individual is unaffected.

Family 13

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	II.1 father	II.2 mother	III.1 proband
D15S128	**0.78**	**22.54**	**6.11**	**5.82**	199-199	203-203	199-203
D15S822	**0.79**	**24.82**	**12.30**	/	286-268	252-286	286-252
D15S219	/	**25.37**	**14.58**	**15.05**	227-223	223-225	227-223
AC024474	/	**26.68**	/	/	205-197	211-205	205-211
D15S1048	**0.84**	**27.44**	**19.12**	**21.07**	221-225	228-228	221-228
D15S165	**0.77**	**28.84**	**20.24**	**22.64**	197-183	199-183	197-199

Table 4.71 Haplotype analysis results on chromosome 15q13 in consanguineous Family 13. The affected child III.1 was heterozygous for all the markers typed, thus excluding linkage to this locus. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65.

Family 14

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.3 proband	II.4 proband
D15S128	**0.78**	**22.54**	**6.11**	**5.82**	208-212	202-206	208-206	212-202
D15S219	/	**25.37**	**14.58**	**15.05**	223-203	227-227	223-227	203-227
AC024474	/	**26.68**	/	/	203-211	203-213	203-213	211-203
D15S1048	**0.84**	**27.44**	**19.12**	**21.07**	220-224	222-224	220-222	224-222
D15S1031	**0.74**	**29.66**	**21.58**	**24.63**	245-249	241-249	245-241	249-241
D15S1010	**0.80**	**30.66**	**23.89**	/	211-227	235-227	211-235	227-235

Table 4.72 Haplotype analysis results on chromosome 15q13 in non-consanguineous Family 14. Patients II.3 and II.4 have inherited different paternal haplotypes, therefore linkage could not be supported. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65.

Family 15

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 proband	II.3 proband	II.4 proband
D15S128	0.78	22.54	6.11	5.82	203- 203	203-193	203- 203	203-193	203- 203	203-193
D15S822	0.79	24.82	12.30	/	282-294	282-298	282-282	294-298	282-282	294-298
D15S219	/	25.37	14.58	15.05	223-225	223-225	223-223	225-225	223-223	225-225
AC024474	/	26.68	/	/	209-209	211-209	209-211	209-209	209-211	209-209
D15S1048	0.84	27.44	19.12	21.07	226-202	202-222	226-202	202-222	226-202	202-222
D15S165	0.77	28.84	20.24	22.64	183-183	187-183	183-187	183-183	183-187	183-183

Table 4.73 Haplotype analysis results on chromosome 15q13 in non-consanguineous Family 15 are not consistent with linkage to this locus. Note how affected fetus II.3 has different haplotypes (highlighted in pink) compared to the other two affected fetuses II.2 and II.4 (highlighted in purple), but the same haplotypes as unaffected sib II.1. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65. Unaff. means individual is unaffected.

Family 16

Markers Chromosome 15	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.2 unaff.	II.3 proband
D15S128	0.78	22.54	6.11	5.82	199-205	197-199	199-197	199-197
D15S822	0.79	24.82	12.30	/	303-259	299-287	303-299	303-299
D15S1048	0.84	27.44	19.12	21.07	219-225	211-219	219-211	219-211
D15S165	0.77	28.84	20.24	22.64	199-183	183-183	199-183	199-183

Table 4.74 Haplotype analysis results on chromosome 15q13 in non-consanguineous Family 16. Patient II.3 has inherited the same pair of haplotypes as her unaffected sib II.2, therefore, linkage could not be supported. Heterozygosity, physical and genetic distances of each marker are indicated. Orange lines define the minimal critical interval determined by the results shown in Table 4.65. Unaff. means individual is unaffected.

Family 15

Markers Chromosome 14	Het	Mb Build 33	cM Marshfield	cM deCode	I.1 father	I.2 mother	II.1 unaff.	II.2 proband	II.3 proband	II.4 proband
D14S739	0.83	80.26	91.62	81.73	180-192	180-188	180-180	180-180	192-180	192-188
D14S68	0.91	86.62	95.89	/	138-138	129-138	138-129	138-129	138-129	138-138
D14S1044	0.67	88.06	99.88	90.62	205-201	199-201	205-199	205-199	201-199	/
D14S617	0.78	90.19	105.53	94.47	158-146	142-146	158-142	158-146	146-142	146-146

Table 4.75 Haplotype analysis results on chromosome 14q24-q32 in non-consanguineous Family 15 are not consistent with linkage to this locus. The three affected fetuses have in fact different haplotypes across the region determined by the results shown in Table 4.19. Heterozygosity, physical and genetic distances of each marker are indicated. Unaff. means individual is unaffected.

4.3.6 Summary

In conclusion, genotyping and haplotype analysis was performed in a total of twelve consanguineous and four non-consanguineous JATD families across the four main putative JATD loci identified on chromosomes 3, 8, 12 and 15.

Five families were consistent with linkage to 3q24-q26. The results in three other families supported linkage to 8q24. Three families were consistent with linkage to 15q13, while none of the families typed were consistent with linkage to the 12q23-q24 locus. Homozygous regions were detected on both chromosomes 3 and 15 in Family 8 and on both chromosomes 8 and 15 in Family 9. The region on 14q24-q32 was investigated only in two families due to time constraints. Each region was refined as tightly as possible by typing a large number of closely spaced microsatellite markers, thus allowing the search for suitable candidate genes to begin.

4.4 Screening of candidate genes

The three putative JATD loci identified in this study (3q24-q26, 8q24 and 14q24-q32) were analysed for candidate genes using three main tools to interrogate human genome data, namely, the Ensembl genome browser, the Golden Path at the UCSC human genome browser and the NCBI Map View (Build 33, April 2003 freeze). These three tools were used to generate an integrated view of genetic and physical maps across the defined loci and to locate genes on the chromosomal assemblies. Candidate genes were then selected in each interval and screened for causative mutations. The nomenclature for describing sequence variants is according to den Dunnen and Antonarakis (2000) and the new recommendations at http://www.hgvs.org/mutnomen/recs.html (den Dunnen and Antonarakis, 2000).

4.4.1 Chromosome 3q24-q26

4.4.1.1 SHOX2 as a candidate gene

The common homozygous region originally identified in affected cousins IV.1 and IV.3 in Family 1 was about 17 cM/23 Mb in size between markers D3S1555 and D3S1574 and contained 145 known or predicted genes. A good candidate gene in this region was thought to be the short stature homeobox 2 gene *SHOX2* (NM_006884.1), located at 3q25-q26.1, between markers D3S1607 and D3S3692 (see Table 4.24). This gene lies in the common homozygous interval in Family 1, 5, 6 and 8.

SHOX2 is closely related to the pseudoautosomal homeobox gene *SHOX,* defects in which are known to lead to idiopathic short stature (Rao *et al.*, 2001), Leri-Weill syndrome (Shears *et al.*, 1998) and short stature associated with Turner syndrome (Clement-Jones *et al.*, 2000). Homozygous *SHOX* mutations have been shown to cause Langer-type mesomelic dwarfism (Shears *et al.*, 2002). *SHOX2* is the othologue of the mouse gene *Og12x*. *SHOX2* is expressed in a variety of human embryonic tissues including limbs, metanephric mesoderm, heart and pharyngeal arches (Clement-Jones *et al.*, 2000). During mouse embryonic development, expression of *Og12x* could be detected in mesoderm derivatives that contribute to bone and cartilage formation in the limb, ribs and spinal cord (Blaschke *et al.*, 1998).

4.4.1.2 Structure of SHOX2

SHOX2 was cloned and mapped to chromosome 3q25-q26 by Blaschke and colleagues (Blaschke *et al.*, 1998). Two different isoforms were identified, *SHOX2a* and *SHOX2b* and their mRNA transcripts were available in the database (NM_006884.1 and NM_003030.2 respectively) providing the full coding sequence. The intron/exon structure was determined by BLAST interrogation of the transcript sequences against the genomic sequence using the BLAST program available at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). The coding sequence was then translated

using Biology Workbench (http://workbench.sdsc.edu) to identify the start codon. The larger transcript *SHOX2a* (NM_006884.1) is comprised of five exons and encodes a protein of 331 amino acids.

Primers to amplify the five exons were designed and synthesised as described in section 2.3.2.3. Primer details and PCR conditions are shown in Table 2.11. The primers were spaced so that splice sites as well as exons could be screened.

4.4.1.3 Mutation screening

PCR amplification of *SHOX2* was carried out as described in section 2.3.2.5. Each exon was successfully amplified by PCR in both the carrier father (III.1) and affected child (IV.1) from Family 1; the carrier mother (I.2) and affected child (II.2) from Family 6; the carrier mother (I.2) and affected child (II.1) from Family 8 as well as in the father (I.1) from Family 5. The patient II.1 in Family 5 could not be screened as there was little DNA and this was difficult to amplify. A single product of the expected size was obtained for each individual exon. Therefore, no evidence of a small homozygous insertion or deletion within any of the exons was detected on PCR in any of the patients, except for the affected child in Family 5 who was not screened.

All five exons were screened for mutations by direct sequencing using protocols detailed in section 2.3.8. Data were analysed using Sequencher 4.1 (see section 2.3.8.3) in both the carrier father (III.1) and affected child (IV.1) from Family 1, the carrier mothers I.2 and I.2 from Family 6 and Family 8 respectively as well as the father (I.1) from Family 5. No pathogenic mutations were identified. Two known intronic variants were found in all the individuals screened: c.557+42G>A (dbSNP4462902) in intron 2 and c.703+131T>C (dbSNP9833383) in intron 4. The affected child IV.1 was homozygous for both changed alleles as were his unaffected father and unaffected individuals I.2 (Family 6) and I.1 (Family 5). The unaffected mother I.2 from Family 8 was heterozygous for both variants.

SHOX2 lies in the homozygous region common to consanguineous Families 1, 5, 6 and 8. However, assuming linkage to 3q24-q26 in non-consanguineous Family 14, *SHOX2*

lies outside the interval identified by the recombination event in unaffected child II.1 from this family (see Table 4.24). Twenty-eight known or predicted genes map to the 8 cM/12 Mb homozygous region common to Family 1, 5, 6, 8 and 14 between markers D3S3575 and D3S3523, although only eighteen were thought to represent likely candidate genes (Table 4.76).

4.4.2 Chromosome 8q24

4.4.2.1 WISP1 as a candidate gene

The common homozygous region originally identified in affected sibs VI.3 and VI.5 in Family 2 was about 13 cM/ 8 Mb in size between markers AC103725 and AC087711-2 (Table 4.14). A schematic representation of the genes in the region is given in Figure 4.15 and their details in Table 4.77. A good candidate gene in the region was thought to be *WISP1* (NM_003882.2) located at 133.9 Mb on 8q24.1-q24.3, between markers D8S529 and D8S256. *WISP1* encodes the Wnt1-inducible signalling pathway protein 1 (WISP1/CCN4) which belongs to the CCN family of secreted proteins (connective tissue growth factors) including cystein-rich 61 (CYR61/CCN1), connective tissue growth factor (CTGF/CCN2), nephroblastoma overexpressed (NOV/CCN3), Wnt1-inducible signalling pathway proteins-2 (WISP2/CCN5) and -3 (WISP3/CCN6). Collectively, these proteins regulate fundamental biological processes, including cell adhesion, growth, differentiation, apoptosis, embryogenesis, vascular diseases tumorigenesis and skeletogenesis.

WISP1 was thought to be a good functional candidate for JATD. Recently, the role of WISP1 during skeletogenesis has been elucidated and expression during embryonic development was shown in osteoblast and osteoblastic progenitor cells of the perichondral mesenchyme (French *et al.*, 2004). In human tissues, *WISP1* expression was seen in a variety of adult organs including kidney, pancreas, lung, heart, ovary and spleen (Pennica *et al.*, 1998). Furthermore, mutations in the *WISP3/CCN6* gene are associated with progressive pseudorheumatoid dysplasia (PPAC; MIM 208230), an

Gene	OMIM	Description of protein
D3S3575		
SCHIP1		Schwannomin-interacting protein 1; interacts with NF2 (neurofibromatosis protein type 2); widely expressed
IL12A	**161560**	Interleukin 12A precursor; subunit of cytokine that acts on T & natural killer cells; important in resistance to TB and leishmania
AB037795		Not RefSeq gene; hypothetical protein, unknown function; WD40 repeats; Aceview says widely expressed
SMC4L1	**605575**	Structural maintenance of chromosome 4; important role in DNA repair/condensation & segregation of mitotic chromosomes; housekeeping gene; widely expressed
TSBF1		Tumour suppressor TSFB1-like; hypothetical protein, unknown function; zinc-binding motifs; not widely expressed
KPNA4	**602970**	Karyopherin alpha 4 (importin alpha 3); mediates nuclear import; housekeeping gene; expressed in testis, ovary, intestine & pancreas
FLJ22595		Hypothetical protein, unknown function; ADP ribosylation factor; not widely expressed, but not expressed in the right tissues
PPM1L		Hypothetical protein; probable protein phosphatase type 2C; not widely expressed - brain, muscle
B3GALT3	**603094**	Beta-1,3-galactosyltransferase subunit 3; housekeeping gene; membrane-bound glycoprotein - role in protein glycosylation; mutations cause globoside-deficient blood group phenotype
CGI-07		CGI-07 protein, possibly related to ribosomal function or nonsense-mediated mRNA decay; very widely expressed
ADMP		Anti-dorsalising morphogenetic protein; TGF-beta homolog related to BMP-3 with important role in dorsal-ventral patterning v. early in development; not widely expressed
BC027488		Not a Refseq gene - hypothetical protein
AF109183		Endogenous retrovirus mRNA
SI	**222900**	Sucrase isomaltase; congenital deficiency causes sucrose intolerance

SLITRK3		SLIT and NTRK-like family, member 3
BCHE	**177400**	Butyrylcholinesterase precursor; mutated in suxamethonium sensitivity (apnea)
FLJ23049		Hypothetical protein, unknown function; not widely expressed - testis, brain, lung
SERPINI2	**605587**	Serine protease inhibitor 14; expressed mainly/exclusively in pancreas - role in growth-suppressing pathway
WD40.166		Not a Refseq gene - hypothetical protein
PDCD10		Programmed cell death 10; resembles proteins that participate in apoptosis; very widely expressed
SERPINI1	**602445**	Protease inhibitor 12 (neuroserpin); expressed mainly in brain; mutated in familial encephalopathy
GOLPH4	**606805**	Golgi phosphoprotein 4; may assist in protein processing & transport in the Golgi; widely expressed
BC011266		Not a Refseq gene - hypothetical protein
EVI1		Ecotropic virus integration 1 site & Myelodysplasia syndrome protein 1; different transcripts - EVI1 & MDS1/EVI1; role as transcriptional activator; involved in fusions in AML; not very widely expressed
MDS1	**600049**	Myelodysplasia syndrome 1
ARPM1		Actin-related protein M1; probable role in cytoskeletal organisation; widely expressed
MYNN	**606042**	Myoneurin; zinc finger transcription factor; mainly in skeletal muscle, also testis, ovary, placenta; clones include cartilage, kidney
MGC27085		Hypothetical protein, unknown function; leucine-rich repeat; quite widely expressed incluing kidney & eye
D3S3523		

Table 4.76 List of genes on chromosome 3q24-q26 between markers D3S3575 and D3S3523. While several of these genes can be ruled out on the grounds of inappropriate expression patterns and/or known roles in other human disorders (highlighted in grey), only eighteen of them were thought to represent possible candidate genes requiring further investigation. Many are genes encoding hypothetical proteins of unknown function.

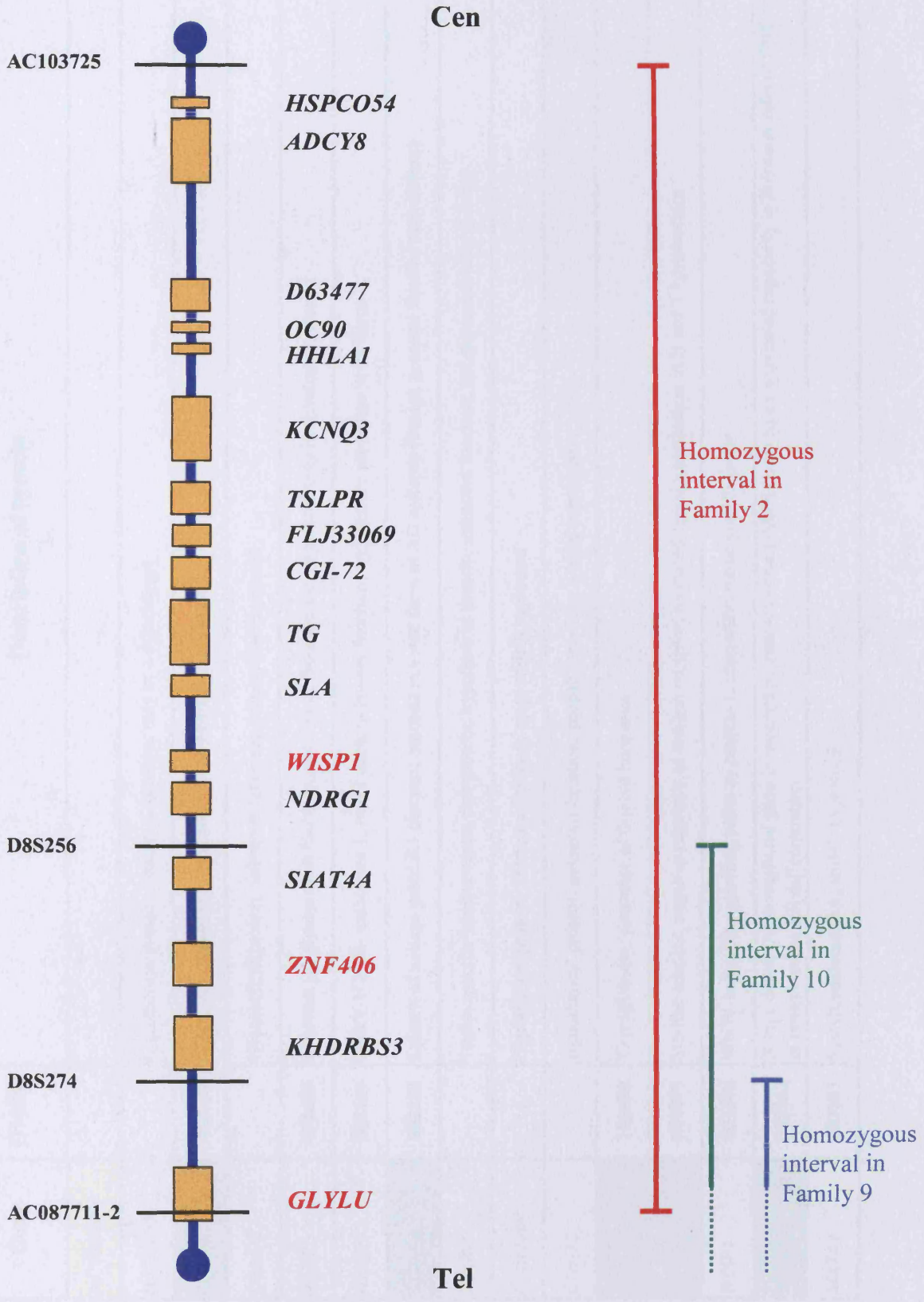


Figure 4.15 Schematic representation of candidate genes lying in the homozygous interval identified in Family 2 on chromosome 8q24 (left). Overlapping regions of homozygosity between Families 2, 9 and 10 are shown on the right. Microstallite markers that define each intervals are also listed. In red are highlighted the genes screened for mutations. For a brief description of each gene see Table 4.77.

Gene	OMIM	Description of protein
AC103725		
HSPC054		Hypothetical protein, unknown function; may be a pseudogene
ADCY8	**103070**	Adenylate cyclase type 8; predominantly expressed in brain
D63477		Hypothetical protein, unknown function; quite widely expressed
OC90	**601658**	Otoconin 90-glycoprotein found in otoconia of the inner ear, expressed only in developing otocyst
HHLA1	**604109**	HERV-H LTR- associated 1; novel gene, unknown function; endogenous retrovirus in an intron
KCNQ3	**602232**	Subunit of voltage gated K+ channel; mutated in a rare form of AD epilepsy (benign neonatal familial convulsions)
TSLRP		Testis- specific leucine-repeat rich protein; hypothetical protein, unknown function; not expressed only in testis
FLJ33069		Hypothetical protein, unknown function; quite widely expressed
CGI-72		Hypothetical protein, unknown function; probably nuclear; widely expressed
TG	**188480**	Thyroglobulin; precursor of thyroid hormones
SLA	**601099**	Src-like adaptor; negative regulator of antigen receptor; mediated signal transduction in B and T lymphocytes
WISP1	**603398**	WNT1-inducible signalling pathway protein 1; connective tissue growth factor
NDRG1	**605262**	N-myc downstream regulated gene 1; mutated in rare recessive HSMN (Lom type); expressed especially in Schwann cells of PNS; role in growth arrest/cell differentiation
SIAT4A	**607187**	Syalyltransferase 4A; widely expressed

ZNF406		Zinc finger protein 406; putative zinc finger transcription factor; widely expressed
KHDRBS3		Testis signal transduction and activation of RNA (T-STAR); expressed mainly in testis, skeletal muscle, brain and heart
GLYLU		Hypothetical protein, unknown function; expressed especially in brain and testis
AC087711-2		

Table 4.77 List of genes on chromosome 8q24 between markers AC103725 and AC087711-2. A few of these genes can be ruled out on the grounds of inappropriate expression patterns and/or known roles in other human disorders (highlighted in grey), whereas others represent possible candidate genes requiring further investigation. Many are genes encoding hypothetical proteins of unknown function. *WISP1* , *GLYLU* and *ZNF406* were screened for mutations in this study.

autosomal recessive skeletal disorder characterized by continued cartilage loss and bone destruction (Hurvitz *et al.*, 1999).

4.4.2.2 Structure of WISP1

WISP1 was cloned and mapped to chromosome 8q24.1-q24.3 by Pennica and colleagues (Pennica *et al.*, 1998). Alternative splicing of the gene generates two transcripts (NM_003882.2 and NM_080838). The larger transcript (NM_003882.2) comprises five exons and encodes the full length protein of 367 amino acids. The transcript sequence, the intron/exon boundaries, as well as the full protein sequence were available in the database at the Ensembl Human Genome Browser.

Primers for the five exons were designed and synthesised as described in section 2.3.2.3. Primer details and PCR conditions are shown in Table 2.12. The primers were spaced so that splice sites as well as exons could be screened.

4.4.2.3 Mutation screening of WISP1

All five exons were mutation screened by direct sequencing using protocols detailed in section 2.3.8 and analysed using Sequencher 4.1 (see section 2.3.8.3) in the mother (V.6) from consanguineous Family 2 as well as in thirteen sporadic affected individuals: probands III.1 from Family 13, II.3 from Family 18, II.1 from Family 19, II.1 from Family 20, II.2 from Family 21, II.1 from Family 24, II.1 from Family 25, II.1 from Family 26, II.1 from Family 27, II.4 from Family 28, II.1 from Family 29, II.3 from Family 30, II.1 from Family 31. DNA samples from the three probands in Families 17, 22 and 23 were not available for screening, therefore the six unaffected parents were screened instead. A control DNA was also included in the analysis.

No pathogenic mutations were identified, although a number of variants, mainly intronic, were found during sequence analysis. A nucleotide substitution was identified

in exon 3 in affected individual II.1 from Family 31 (c.550G>A) resulting in a p.Val184Ile alteration.

In exon 5, a silent substitution c.921T>C, which did not change Asn at position 307, was identified. Six probands were heterozygous for the change, as was the control. This synonymous variant was listed in the NCBI SNP database (dbSNP3739261).

In the proband from Family 24, a heterozygous G>A variant at –115 from the ATG was found in the 5'UTR. It was not listed in the NCBI SNP database as a known SNP.

A heterozygous c.70-39T>C intronic variant (dbSNP10089461) was identified in intron 1 in five affected individuals (Families 13, 19, 24, 25 and 28) and in the unaffected parents from Families 2, 17, 22 and 23. Probands in Families 18, 21, 26 and 27 were homozygous for the C allele.

A heterozygous c.611-55G>C intronic variant (dbSNP2280834) was found in intron 3 in almost all the individuals screened, including the control. Exceptions were individuals from Families 17, 18 and 27 who were homozygous for the C allele and probands from Families 29 and 31 who were homozygous for the G allele.

4.4.2.4 GLYLU as a candidate gene

Following genotyping analysis of other JATD families, Families 9 and 10 were consistent with linkage to 8q24. Assuming linkage in these families, the common homozygous interval in Families 2, 9 and 10 was narrowed down to a 2 cM (1.65 Mb) region defined by marker D8S274 at the centromeric end and marker AC087711-2 at the telomeric end (Fig. 4.15). *WISP1* lies outside this interval. Analysis of the region using the UCSC Genome Browser showed no gaps in the sequence and led to the identification of only one gene predicted by Ensembl and Acembly between markers AFM316xb1 and AC087711-2. The gene encoding hypothetical protein LOC51059 (NM_015912.2) was called *GLYLU* by Acembly and was predicted to have four different spliced forms. Supporting evidence for this gene came from Twinscan and SGP gene prediction programs using mouse/human homology, the presence of human

mRNA transcripts (the longest being AK091433) and spliced ESTs. Furthermore, the gene appeared to be conserved in mouse and fugu.

GLYLU was considered an excellent positional candidate, although information about the protein's (hypothetical protein LOC51059) expression and function were poor. It does not belong to any known protein family. Supporting clones came from brain, testis and pooled germ cell tumours. PSORT predicts that it localizes to the nucleus; PROSITE predicts a thiol protease site in exon 12.

4.4.2.5 Structure of GLYLU

Ensembl, Acembly and Twinscan made slightly different predictions of the gene. All exons identified by these programs were analysed and the complete cDNA was assembled and translated using Biology Workbench, the start codon in exon 2 and the longest open reading frame determined. Primers were designed for each exon and synthesised as described in section 2.3.2.3. Primer details and PCR conditions are shown in Table 2.13. The primers were spaced so that splice sites as well as exons could be screened. This novel gene was determined to comprise twenty-two exons.

4.4.2.6 Mutation screening of GLYLU

Mutation screening was carried out in a single parent from each of the three families (Families 2, 9 and 10) consistent with linkage to 8q24. All twenty-two exons were mutation screened by direct sequencing using protocols detailed in section 2.3.8 and analysed using Sequencher 4.1 (see section 2.3.8.3). Each exon was successfully amplified by PCR in the three parents and also in affected individual (VI.3) from Family 2, thus excluding a microdeletion in any of the exons in this family.

No pathogenic mutations were identified, although a number of heterozygous synonymous alterations (that did not alter amino acids, due to third base wobble) were found throughout the gene. Only one nucleotide substitution identified in all three

individuals screened altered an amino acid in exon 8 (p.Arg362Ser). Exon 8 was then sequenced in proband VI.3 from Family 2 and she was homozygous for the serine residue, but so was her unaffected mother. Moreover, exon 8 was thought not to represent a true exon as it was predicted only by Ensembl (transcript ENST00000303060) and was not supported by evidence from any other source.

4.4.2.7 Further genes on 8q24

No further genes were identified in the common region of homozygosity in Families 2, 9 and 10 therefore the attention shifted to the region of overlap between Families 2 and 10 defined by markers D8S256 and D8S274 (Fig. 4.36). A reason for that was the uncertainty about the diagnosis in the affected child from Family 9. He had in fact extremely mild radiological findings and was described as not a classic JATD case. It was thought that the homozygosity across this region might have been a chance finding considering the fact that he also showed homozygosity across the 15q13 locus.

Three genes lie in this 3 Mb interval with no gaps in the sequence, *SIAT4A*, *ZNF406* and *KHDRBS3* (Fig. 4.15). *SIAT4A* (sialyltransferase 4A) is a glycosyltransferase widely expressed in adult tissues, cancer and fetal tissues. *ZNF406* contains multiple zinc finger motifs and an AT-hook and appears to encode a zinc finger transcription factor. It is expressed in a wide variety of adult tissues and tumours and was predicted by PSORT to localise to the nucleus. *KHDRBS3* alternatively named *T-STAR* (testis-signal transduction and activation of RNA) is a Sam68-like phosphotyrosine nuclear protein involved in RNA processing and signal transduction, specifically in regulating alternative splicing in response to extra-cellular signals. It is expressed in testis, skeletal muscle, brain and heart. *ZNF406* was thought to be the best candidate gene of the three, and was therefore sequenced first.

4.4.2.8 Structure of ZNF406

ZNF406 (NM_020863.2) was predicted by Acembly, Ensembl, Twinscan and SGP gene prediction programs. Ensembl predicts three slightly different transcripts. Exon 1 (transcript ENST318135) seems to contain the true start codon of the gene encoding a methionine, preceded by a convincing Kozak sequence and a CpG island. The complete cDNA was assembled and translated using Biology Workbench. The larger transcript comprises seventeen exons and encodes a protein of 1284 amino acids. The protein appears to be conserved not only in mouse, but also partially in fugu. Primers were designed for each exon and synthesised as described in section 2.3.2.3. Five primer pairs were designed to screen exon 7. Primer details and PCR conditions are shown in Table 2.14. The primers were spaced so that splice sites as well as exons could be screened.

4.4.2.9 Mutation screening of ZNF406

All seventeen exons were screened for mutations by direct sequencing using protocols detailed in section 2.3.8 and analysed using Sequencher 4.1 (see section 2.3.8.3) in the unaffected father (V.5) from Family 2 and the unaffected mother (I.2) from Family 10.

In the mother from Family 10, two intronic variants and four heterozygous synonymous polymorphisms were found throughout the gene. A heterozygous c.20-41A>G variant was identified in intron 1. A heterozygous c.637+74A>G intronic known variant (dbSNP2277138) was found in intron 5. A heterozygous synonymous c.990T>C known alteration (dbSNP3739422) that did not change the Tyr residue at position 371 was identified in exon 7. In exon 8 a c.2409A>G reported alteration (dbSNP894344) which did not change the Asp residue 844 was found. In exon 17 a non-reported heterozygous synonymous variant (c.3621G>A) resulted in the same amino acid residue (Thr 1248).

A poor quality sequence was obtained for exon 9. Primers ZNF4069F and ZNF4069R (listed in Table 2.14) were therefore used to amplify a 426 bp PCR fragment containing exon 9 of the gene and PCR products from both parents (Family 2 and 10) were

subcloned using the PCR-Script™ Amp Cloning kit as described in section 2.3.9. Colony PCR was then performed as explained in section 2.3.10 and the products were sequenced as described in section 2.3.8. In individual I.2 from Family 10 a heterozygous synonymous known variant c.2493T>C (dbSNP3739425) that did not change the Ser residue 872 was identified in exon 9.

In the father (V.5) of Family 2, a heterozygous synonymous known alteration c.1257T>C (dbSNP3739423) that did not change the Arg residue 460 was identified in exon 7. In exon 16 a synonymous reported variant c.3384C>T (dbSNP11778717) resulted in the same amino acid residue (Ala 1169). In this family, two non-synonymous alterations were identified in exons 3 and 4.

In exon 3, a heterozygous c.190G>A transition resulting in an amino acid change (p.Gly105Arg) was identified in the father (V.5). This is not a conservative change and the mouse protein has Gly at the same position. To analyse the segregation of the alteration with the disease within the family, a *HpaII* restriction digest was performed (see section 2.2.4). Primers ZNF4063F and ZNF4063R (listed in Table 2.14) were used to amplify a 387 bp PCR fragment containing exon 3 of the gene. A *HpaII* restriction site occurs after base 292 of this fragment if the nucleotide G is present. After digestion, normal 387 bp, 292 bp and 95 bp bands were detected on a 4% agarose gel in the unaffected father (V.5), mother (V.6) and unaffected sib (VI.2) indicating that they were heterozygous for the alteration (Fig. 4.16). In unaffected child VI.4 only the 292 bp and 95 bp bands were detected on agarose gel, indicating that she was homozygous for the G allele. In the affected children VI.3 and VI.5 only the uncut 387 bp PCR product was detected, indicating that they were both homozygous for the variant allele A and therefore homozygous for the Arg residue at position 105. This amino acid substitution does not occur in a region of the gene of obvious functional significance and Blast interrogation revealed the presence of one human mRNA IMAGE clone (3840083) which has the same base alteration.

In exon 4, a heterozygous c.304C>T substitution results in a p.Pro143Ser change. This is not a conservative change and the mouse protein also has a Pro at the same position, although this was a reported variant (dbSNP12541381). To analyse the segregation of the alteration with the disease within the family, an *AciI* restriction digest was

Figure 4.16 Segregation of the c.190G>A transition in Family 2. The alteration destroys a *HpaII* site that occurs after base 292 of the 387 bp PCR product ZNF4063F/R containing exon 3. In affected individuals VI.3 and VI.5, PCR amplification of genomic DNA using primers pair ZNF4063F and ZNF4063R, followed by digestion with *HpaII* and analysis on a 4% agarose gel, gave the 387 bp fragment representing the uncut PCR product indicating that they were homozygous for the variant allele A. In unaffected individuals V.5, V.6 and VI.2 all three bands were seen, indicating that they were heterozygous for the change. In unaffected individual VI.4 only the 292 bp and the 95 bp fragments were detected, indicating that she was homozygous for allele G. Solid symbols represent affected individuals. Lane 2, V.5. Lane 3, V.6. Lane 4, VI.2, Lane 5, VI.4. Lane 6, VI.3. Lane 7, VI.5. The size marker in Lanes 1 and 8 is a 50 bp DNA ladder.

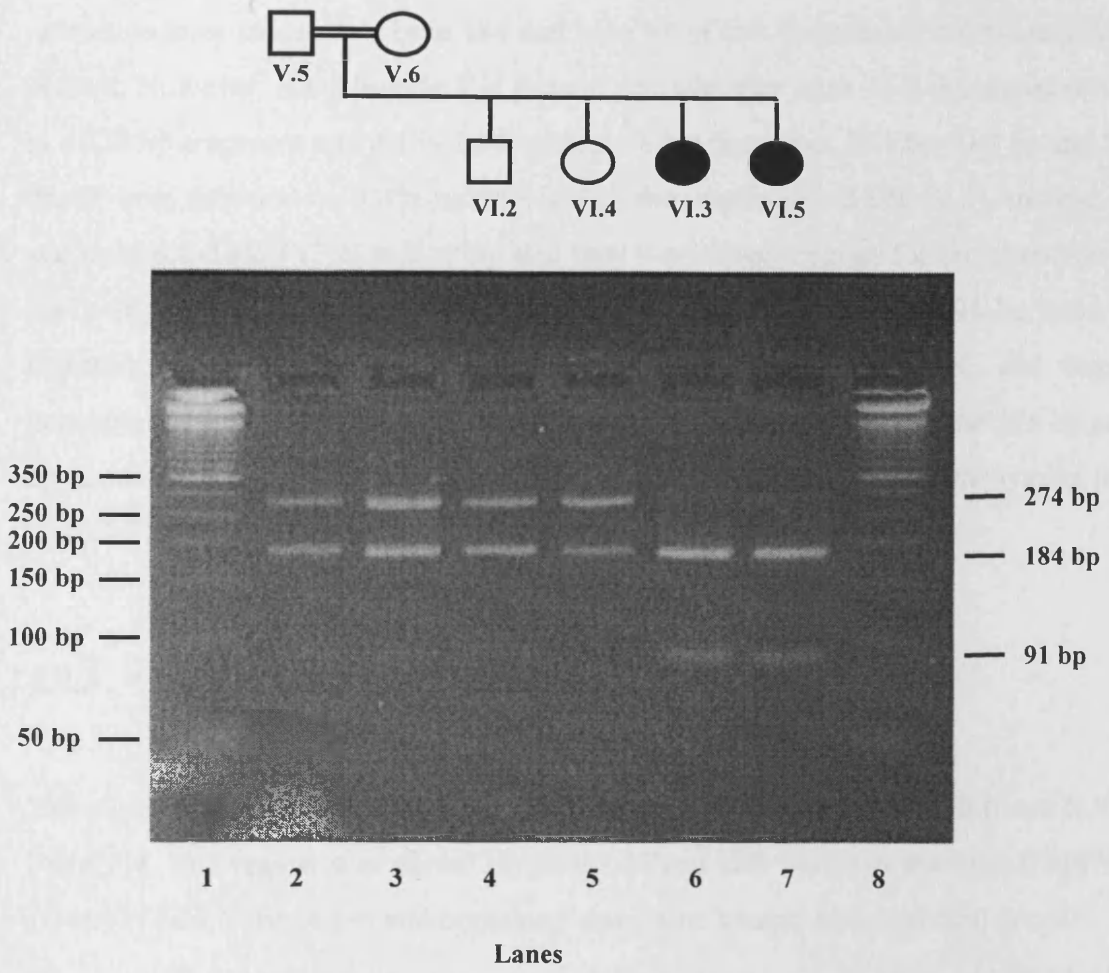


Figure 4.17 Segregation of the c.304C>T transition in Family 2. *AciI* cuts the 459 bp PCR product ZNF4064F/R (containing exon 4) in two positions, at 184 and 91. The alteration destroys the *AciI* site that occurs after base 91 of the PCR product. In affected individuals VI.3 and VI.5, PCR amplification of genomic DNA using primers pair ZNF4064F and ZNF4064R, followed by digestion with *AciI* and analysis on a 4% agarose gel, gave 184 bp and 91 bp fragments, indicating that they were homozygous for allele C. In unaffected individuals V.5, V.6 and VI.2 all three bands were seen, indicating that they were heterozygous for the change. In unaffected individual VI.4 only the 274 bp and the 184 bp fragments were detected, indicating that she was homozygous for allele T. Solid symbols represent affected individuals. Lane 2, V.5. Lane 3, V.6. Lane 4, VI.2, Lane 5, VI.4. Lane 6, VI.3. Lane 7, VI.5. The size marker in Lanes 1 and 8 is a 50 bp DNA ladder.

performed (see section 2.3.4). Primers ZNF4064F and ZNF4064R (listed in Table 2.14) were used to amplify a 459 bp PCR fragment containing exon 4 of the gene. Two *AciI* restriction sites occur after base 184 and base 91 of this fragment if the nucleotide C is present. However, if nucleotide T is present, the site after base 91 is destroyed resulting in a 274 bp fragment and a 184 bp fragment. After digestion, 274 bp, 184 bp and 91 bp bands were detected on a 4% agarose gel in the unaffected father (V.5), mother (V.6) and unaffected sib (VI.2) indicating that they were heterozygous for the alteration (Fig. 4.17). In the affected children VI.3 and VI.5 a 184 bp band and a 91 bp band were detected, indicating that they were both homozygous for allele C and therefore homozygous for the Pro at residue 143. In unaffected child VI.4 only the 274 bp and 91 bp bands were detected on an agarose gel, indicating that she was homozygous for the T allele. Therefore, it is very unlikely that the change is pathogenic.

4.4.3 Chromosome 14q24-q32

The region on chromosome 14q24-q32 was homozygous in probands II.6 and II.8 from Family 4. The region was about 19 cM/14 Mb in size between markers D14S74 and D14S977 (see Table 4.19) and contained sixty-nine known and predicted genes.

Interestingly, *BBS8 (TTC8)*, mutations in which are responsible for Bardet-Biedl syndrome (BBS; MIM 209900), lies in this interval (Ansley *et al.*, 2003). BBS is a genetically heterogeneous disease with linkage to eight loci characterized by multiple clinical features and a complex model of inheritance (Ansley *et al.*, 2003; Katsanis *et al.*, 2001). Six BBS genes have been identified and recent studies have suggested that ciliary defects compromising intraflagellar transport (IFT) are responsible, at least in part, for the BBS phenotype (Kim *et al.*, 2004; Kulaga *et al.*, 2004; Blacque *et al.*, 2004). Three proteins, BBS4, BBS5 and BBS8 localized to ciliated borders in tissues, and to the basal body in cultured cells. The latter have a role in microtubule organization of ciliated cells (Kim *et al.*, 2004; Ansley *et al.*, 2003; Badano *et al.*, 2003). BBS4 and BBS8 contain several tetratricopeptide repeats (TPRs) that are

involved in protein-protein interactions (Ansley *et al.*, 2003). BBS8 is a centrosomal protein and localizes specifically to ciliated structures, such as the connecting cilium of the retina and columnar epithelial cells in the lung (Ansley *et al.*, 2003). BBS4 is a pericentriolar protein necessary for recruiting cargo to centriolar satellites by acting as a dynein-associated adaptor protein (Kim *et al.*, 2004).

As described in more detail in section 5.4.3, it has been suggested that genes encoding intraflagellar transport (IFT) proteins might be promising candidate disease genes for JATD (Pazour and Rosenbaum, 2002; Rosenbaum and Witman, 2002). Typical features of JATD such as cystic kidneys and retinal degeneration in combination with skeletal defects and other abnormalities might be caused by defects in cilia, as has been demonstrated for BBS and Senior-Loken syndrome (MIM 266900) (Pazour and Rosenbaum, 2002; Rosenbaum and Witman, 2002; Schuermann *et al.*, 2002).

4.4.3.1 TTC7B as a candidate gene

A gene encoding a tetratricopeptide repeat domain 7B protein *TTC7B* (XM_085175.4) was identified on 14q24-q32 at 89 Mb between markers D14S1044 and D14S995 (see Table 4.19) and represented an excellent candidate gene for JATD. TTC7B contains five TPR domains, one coatomer WD associated region and a domain similar to the prokaryotic *pilF* domain involved in pilus assembly and intracellular trafficking. It is conserved in mouse and fugu and the homologous mouse protein is expressed in bone, kidney, pancreas, retina and lung.

4.4.3.2 Structure of TTC7B

The structure of *TTC7B* (XM_085175.4) was retrieved from the NCBI database and intron-exon boundaries determined from AL832848 (except for exon 1 which was identified in genomic contig NT_026437.1447). Ensembl also predicted exon 3-21 (ENS267557). The complete cDNA was assembled and translated using Biology Workbench. The larger transcript comprises twenty-one exons and encodes a protein of

860 amino acid. Primers were designed for each exon and synthesised as described in section 2.3.2.3. Primer details and PCR conditions are shown in Table 2.15. The primers were spaced so that splice sites as well as exons could be screened.

4.4.3.3 Mutation screening of TTC7B

Seventeen exons were successfully amplified and screened for mutations by direct sequencing using protocols detailed in section 2.3.8 and analysed using Sequencher 4.1 (see section 2.3.8.3) in the affected child II.6 from Family 4. Exons 1B, 2, 8 and 21 could not be amplified and therefore were not screened.

No pathogenic mutations were identified in the coding sequence of the gene. Only a known synonymous variant c.1292A>G (dbSNP342660), which did not change the Pro residue 428 was identified in exon 11.

4.4.4 Further investigations of candidate loci

A yeast two hybrid (Y2H) screen for proteins that might interact with EVC was carried out by Prof. Goodship's group. EVC Y2H positives were checked to see if any of them mapped to the five JATD loci, but none did (personal communication).

Nucleotide and protein homology searches for genes and proteins similar to *EVC* and *EVC2* within the putative JATD loci have been performed in this study with no success.

Chapter 5 JATD Discussion

This study aimed to map the gene or genes for JATD by linkage analysis using the technique of homozygosity mapping. At the beginning of the project, two candidate regions, the *EVC/EVC2* region and the 12p11-p12 region had been proposed for JATD. A mouse model (*srt*) for JATD had also been suggested. *Srt* maps to mouse chromosome 17 and in humans the three regions which are likely to contain the human homologue of the gene at the *srt* locus are on 6p21, 6q25-q27 and 16p13.3. Linkage to all loci was excluded in a number of JATD families.

Following a genome-wide screen in three consanguineous JATD families, three regions were identified, one in each of the families studied. In each case the potential JATD locus was refined by typing other consanguineous and non-consanguineous families. Candidate genes were then chosen in these loci and screened for causative mutations by direct sequencing of coding regions. No mutations have been identified in the genes screened.

The project started when two large consanguineous JATD families (Families 1 and 2) were recruited. In Family 1, where the two affected children are double first cousins, it was calculated that, if DNA could be obtained from the 2 unaffected children in generation IV as well (giving a total of 12 informative meioses), a lod score of 3.6 could be obtained with a highly-polymorphic marker tightly linked to the disease locus. In Family 2, where the two affected sisters are the children of a third-cousin marriage, it was calculated that, if DNA could be obtained from all 3 unaffected siblings (giving a total of 16 informative meioses), a lod score of 4.8 could be obtained. Subsequently, DNA was obtained from a further 10 consanguineous families and 19 non-consanguineous families. A second genome-wide screen was performed in the affected individual in consanguineous Family 4.

5.1 Exclusion of candidate genes/loci

Linkage to the *EVC/EVC2* region, previously thought to be a candidate locus for JATD, was excluded in this study in seven JATD families. These results support those of Krakow *et al.*, who excluded linkage to *EVC* in one JATD family (Krakow *et al.*, 2000). Family 12 was not consistent with linkage to any of the identified JATD loci and a region of homozygosity was identified in the affected fetus across the *EVC/EVC2* locus. The affected fetus was said to have a phenotype compatible both with JATD and with EVC.

Linkage to the candidate JATD locus on 12p11.21-p12.2 in Families 1 and 2 was also excluded. Finally, linkage in Families 1 and 2 to 6p21, 6q25-q27 and 16p13, which are homologous to part of the critical region on mouse chromosome 17 where the *srt* mouse maps (Herron *et al.*, 2002), was also excluded.

5.2 Homozygosity mapping

The homozygosity mapping approach was then adopted. Initially a genome-wide screen in the four affected children from Families 1 and 2 was performed. Calculations suggest that, if the parents are first cousins, the region of homozygosity around the disease locus in the affected child is likely to be about 28 cM; if the parents are second cousins, about 22 cM; if third cousins, smaller still (Genin *et al.*, 1998). It has also been calculated that highly-polymorphic markers spaced on average every 5 cM should be sufficient to demonstrate linkage in four affected individuals who are the offspring of different first or second cousin marriages (Lander and Botstein, 1987). Therefore, it was estimated that the new Single Chromosome Scan linkage marker set from Research Genetics, which has an average marker spacing of 7 cM, would be sufficient; but as explained in section 4.2.1.2, the first-pass screen with this marker set did not give adequate coverage of the genome. Furthermore, only four markers were found at which all four affected children were homozygous: D4S3249 and D4S2394 on different parts

of 4q, D8S1119 on 8q and D17S1298 on 17p. All four regions were excluded when adjacent markers were typed. Therefore a second-pass screen was carried out, using additional polymorphic markers, aiming to fill in any gaps greater than 7-10 cM between the markers tested.

After the second-pass screen, more regions in which all four affected children were homozygous were identified, but subsequently excluded when adjacent markers were found to be heterozygous in the patients. A region that was IBD in all four children was not identified. Inividual regions of homozygosity within each family were then sought, and each investigated further. Two distinct loci were identified in Families 1 and 2 as a result of the second-pass complete genome-wide. In this study, a third candidate locus was identified following another genome-wide screen in Family 4. Two different loci were independently identified by our collaborators in Paris and Birmingham following genome-wide screens. Locus heterogeneity for JATD was therefore demonstrated, with at least five loci identified so far. This had been considered, but not assumed, during the first stages of the project. Therefore, the initial interpretation of the results from the genome-wide screen was misleading as only common regions of homozygosity in both families were investigated.

5.3 Candidate regions

The five JATD loci are on chromosomes 3q24-q26, 8q24 and 14q24-q32, identified in this project and 12q23-q24 and 15q13 as identified by the Paris and Birmingham groups respectively.

5.3.1 Chromosome 3q24-q26

The first locus on 3q24-q26 was identified in Family 1 following a genome-wide screen and was the only large region in the whole genome in which the two affected children were both homozygous for the same alleles. Evidence for linkage to this locus was

supported by genotyping results in three other consanguineous families (Families 5, 6 and 8) and one non-consanguineous family (Family 14). Furthermore, Pakistani Families 1, 5, 8 and 14 shared identical alleles at markers D3S3579, D3S1268 and D3S3668 within the minimal critical interval suggesting the presence of a shared ancestral chromosomal segment inherited from a distant common ancestor and supporting evidence for linkage to this locus.

However, Family 8 was also consistent with linkage to the region on chromosome 15q13, whereas Family 5 was not screened across the 8q24 and 12q23-q24 loci as there was insufficient DNA from the proband. Moreover, a genome-wide screen was not performed in Families 5 and 14 and none of the families were typed across the 14q24-q32 locus as this region was identified in the last stages of the project.

5.3.2 Chromosome 8q24

The second locus was identified on 8q24 in Family 2 following a genome-wide screen. Two more consanguineous families (Families 9 and 10) were found to be consistent with linkage to this locus and a common 2 cM/1.65 Mb region of homozygosity was detected between marker D8S274 at the centromeric end and marker AC087711-2 at the telomeric end. Little DNA was obtained from the mother of affected child II.1 in Family 9 and therefore she could not be typed for all the markers. Furthermore, not all markers typed in this region were fully informative and the homozygosity in the child could be IBS. Doubts were cast about whether the diagnosis of JATD in Family 10 was correct. The affected child II.10 in this family, in fact, also had severe mental retardation, never previously associated with JATD as well as Dandy-Walker

malformations, not a common feature of JATD. The affected child in Family 9 was also described as not being a classical JATD case. Furthermore, he was homozygous for a small region of about 6 cM/2.5 Mb on 15q13 as described by Morgan and colleagues (proband D) as well as the region on chromosome 8 (Morgan *et al.*, 2003).

A genome-wide screen was performed in families 2 and 9 and although 8q24 was the most promising region throughout the genome, other regions might need to be investigated further. In Family 2, for example, a number of quite large gaps from the genome-wide screen are still present and there might well be another homozygous region in the two affected sibs that has not yet been identified.

5.3.3 Chromosome 12q23-q24

The locus on chromosome 12q23-q24 was mapped by collaborator Dr Valerie Cormier-Daire and was supported by data from two JATD consanguineous families. The minimal critical interval of 2 cM/ 3 Mb in size was identified between markers D12S1342 and D12S1339. None of the families analysed in this study were found to be consistent with linkage to this locus.

5.3.4 Chromosome 15q13

The locus on chromosome 15q13 was identified by collaborator Dr Colin Johnson and is supported by data from five consanguineous families and three non-consanguineous families with a maximum two-point Lod score of 3.77 at θ=0 obtained at D15D1031 (Morgan *et al.*, 2003). The minimal critical interval was 1.2 cM/ 1.5 Mb in size and contained 16 known and predicted genes (Morgan *et al.*, 2003). Two of these, *FMN-1* and *GREMLIN*, initially appeared to be good candidates, but none of the linked families were found to harbour mutations in the coding sequence of either gene.

In this study, evidence for linkage to this locus was confirmed in consanguineous Families 3 and 8 ((Morgan *et al.*, 2003), probands A and C respectively). Further typing in these two families refined the boundaries of the region between AC024474 and D15S976. In light of these results, both *FMN-1* and *GREMLIN* lie outside the minimal critical interval and the homozygous region in patient II.1 from Family 3 (proband A) do not overlap with the homozygous region in patient II.1 from Family 9 (proband D). It is still possible that there is a small overlapping homozygous interval of about 300 Kb in the two probands between markers D15S165 and D15S976, although this family might not be linked to this locus. The homozygous region in proband II.1 from Family 9 (about 6 cM/2.5 Mb) was in fact thought to be small considering that the parents were first cousins. Moreover, as the heterozygosity of the markers typed by Morgan and collaborators was low, the father was also homozygous for D15S1013, while marker D15S231 was uninformative as both parents were homozygous (Morgan *et al.*, 2003). Furthermore, as explained in section 5.3.2, Family 9 was also consistent with linkage to the locus on 8q24. The affected child II.1 showed homozygosity through a larger region (16 cM/8 Mb) on chromosome 8q24 than on 15q13, as expected from a first cousin mating.

As explained in section 5.3.1, Family 8 also demonstrated homozygosity across two JATD loci, 3q24-q26 as well as 15q13. It is possible that one of the two regions is IBS and not IBD in the affected child, but thus far this cannot be corroborated because a number of markers on 15q13 were uninformative. Therefore, this family needs to be screened for mutations in candidate genes from both regions. It is possible that once the JATD genes are identified, this family may harbour mutations in two different genes, thus suggesting a more complex mode of inheritance (see section 5.5).

In conclusion, the results described in this project supersede the ones previously reported by Morgan and collaborators (Morgan *et al.*, 2003). The minimal critical interval is located more centromeric than previously thought and does not include *FMN-1* or *GREMLIN*. Families 8 and 9 seem to map also to two loci identified in this study, 3q24-q26 and 8q24 respectively.

5.3.5 Chromosome 14q24-q32

The locus on 14q24-q32 was identified in Family 4 following a genome-wide screen in one of the two affected children. Although most of the markers were uninformative across the locus and therefore more genotyping needs to be carried out, this was the most promising region found in this particular family. Due to time constraints, only one other family (non-consanguineous Family 15) was genotyped across 14q24-q32, and linkage was excluded.

The phenotype in the probands from Family 4 was thought to be atypical for JATD and could be the reason why this family was not consistent with linkage to any other putative JATD loci. The child had postaxial polydactyly in one foot, with dysplastic toenails bilaterally and short third toes bilaterally. He had normal fingernails. Hypoplastic/dysplastic nails are a common feature in EVC and do not occur at all in JATD. According to Dr Brueton's review, postaxial polydactyly is always present in EVC, but only occasionally present in JATD (Brueton *et al.*, 1990). Moreover, the pattern of polydactyly is different. In EVC it is always present in the hands and only rarely in the feet, whereas in JATD when it is present, it usually affects both the hands and feet. Therefore the phenotype in this patient does not fit with either JATD or EVC. Linkage to the *EVC/EVC2* locus was excluded in this family.

5.3.6 Summary of genotyping results

A summary of the genotyping screen carried out in this project is shown in Figure 5.1.

Families 3, 5, 6, 8 and 14 were consistent with linkage to 3q24-q26; families 2, 9 and 10 were consistent with linkage to 8q24; families 3, 8 and 9 were consistent with linkage to 15q13. Homozygous regions were detected on both chromosomes 3 and 15 in Family 8 and on both chromosomes 8 and 15 in Family 9. None of the families were screened across the 14q24-q32 locus identified in Family 4 except for non-

Family	Ethnic origin	Consistent with linkage to: chr 3	chr 8	chr 12	chr15	chr 14	EVC
1	Pakistani	Y	N	N	N	/	N
2	Dutch	N	Y	N	N	/	N
3	Pakistani	N	N	N	Y	/	N
4	Arab	possible	N	N	N	Y	N
5	Pakistani	Y	/	/	N	/	N
6	Turkish	Y	N	N	N	/	/
7	Turkish	possible	N	N	possible	/	/
8	Pakistani	Y	N	N	Y	/	N
9	Italian	N	Y	N	Y	/	N
10	Maroccan	N	Y	N	N	/	?
11	Turkish	N	N	N	N	/	N
12	English	N	N	N	N	/	EVC
13	Italian	N	N	N	N	/	N
14	Pakistani	Y	N	N	N	/	/
15	English	N	N	N	N	N	N
16	Lithuanian	possible	N	possible	N	/	N

Figure 5.1 Summary of genotyping results in JATD families. A genomewide screen was carried out in Families 1, 2, 3 and 4 in this study and in Families 8, 9 and 16 by our collaborators. Y highlighted in orange indicates that the family is consistent with linkage to that locus. N indicates that the family is not consistent with linkage to that locus. / indicates that the region has not been tested. Family 8 was consistent with linkage to both chromosomes 3 and 15, while Family 9 to both chromosomes 8 and 15. Family 12 was found to be consistent with linkage to the *EVC/EVC2* locus, while Family 11 was found to have a different condition.

consanguineous Family 15 for which linkage to this locus could not be demonstrated. Linkage to the *EVC/EVC2* locus was not excluded in all the families. Family 5 could not be typed across the chromosome 8 or 12 regions as little DNA from the proband was available. In Family 7, as DNA from the patient was not available, haplotype analysis was carried out only in the parents and the two unaffected sibs and could be consistent with linkage to both regions on chromosome 3 and 15. Non-consanguineous Family 16 could also be consistent with linkage to both 3q24-q26 and 12q23-q24.

The parents of patient III.1 in Family 13 were thought to be distantly related. Assuming this, the affected child was not consistent with linkage to any of the JATD candidate loci, even though this family was not typed across the 14q24-q32 region. In non-consanguineous Family 15 linkage could not be demonstrated to any of the five loci or the *EVC/EVC2* locus, suggesting that there is at least one additional locus to be identified.

Following screening of the candidate regions, Family 11 and 12 were diagnosed with different conditions. Family 11 was not consistent with linkage to any of the JATD putative loci or the *EVC/EVC2* locus. An exact diagnosis could not be determined in the affected child. However, Prof. Christine Hall suggested acromesomelic dysplasia, Maroteaux type after reviewing X-rays from the patient. Dr Valerie Cormier-Daire suggested an atypical form of EVC. Family 12 on the other hand was consistent with linkage to the *EVC/EVC2* locus, therefore DNA samples were sent to Prof. Judith Goodship to investigate further the presence of a mutation in *EVC* or *EVC2* as the possible cause of the phenotype in this family. This investigation is ongoing. This highlights the necessity of formulating a correct diagnosis in each family before the commencement of any linkage study to avoid confusion in the interpretation of the results.

A formal genotype phenotype study is not possible with a small number of families such as these, however it is interesting to note that both severely and mildly affected families map to the same regions and may suggest that phenotypic variation in JATD reflects allelic heterogeneity and not locus heterogeneity. A consistent relationship between families linked to a particular locus and their countries of origin could not be shown, although on 3q24-q26 four families out of five were Pakistani.

The region on chromosome 3q24-q26 was the first one to be identified and was supported by data from five families. Linkage to the 8q24 region was supported by data from three families, although in Families 9 and 10 doubts arose about the diagnosis. The region on 14q24-q32 was the last one to be identified and requires further investigation. Nevertheless, the search for candidate genes in these three regions has begun.

5.4 Candidate genes

5.4.1 Chromosome 3q24-q26: *SHOX2*

The short stature homeobox 2 gene *SHOX2* was originally thought to be a good candidate gene for JATD. *SHOX2* maps to 3q25-q26.1 and lies within the common homozygous interval in patients from consanguineous Families 1, 5, 6 and 8 between markers D3S1607 and D13S3692 (see Table 4.24). The two affected sibs from non-consanguineous Family 14 also have the same haplotypes in the interval (see Table 4.23).

SHOX2 belongs to the homeobox family of genes, which encode proteins characterized by a 60 amino acid DNA binding domain, and are known to be involved in pattern formation during embryonic development. Mutations in homeobox genes have been described in a number of human disorders: *HOXD13* in synpolydactyly (SPD; MIM 186000); *HOXA13* in hand-foot-genital syndrome (HFG; MIM 140000); *PAX3* in Waardenburg syndrome (WS1; MIM 193500); *PAX6* in aniridia (AN2; MIM 106210); *HLXB9* in Currino-Triad syndrome (MIM 146450). The closely related pseudoautosomal homeobox gene *SHOX* is known to be mutated in Leri-Weill syndrome, a mesomelic short stature syndrome and homozygous *SHOX* mutations have been shown to cause Langer type mesomelic dwarfism (Shears *et al.*, 2002; Shears *et al.*, 1998). In addition, defects in *SHOX* are known to lead to idiopathic short stature and short stature associated with Turner syndrome (Clement-Jones *et al.*, 2000).

Two isoforms, SHOX2a and SHOX2b, have been identified. The *SHOX2b* transcript differs from the *SHOX2a* transcript in that the translation start site lies 131 codons downstream from that of *SHOX2a*. *SHOX2b* is also lacking an exon in the C-terminal region. Expression studies in human embryos have shown that *SHOX2* is expressed in a variety of organ systems, including limbs, pharyngeal arches and genital tubercles (Clement-Jones *et al.*, 2000). *SHOX2* is also expressed in heart, liver, lung, and pancreas (Blaschke *et al.*, 1998). The murine orthologue *Og12x* shows a high degree of conservation with *SHOX2* and is expressed in the fore- and hind-limb buds, the ribs, the metanephric mesoderm as well as the central nervous system, the heart and the craniofacial tissues during mouse embryonic development (Blaschke *et al.*, 1998; Clement-Jones *et al.*, 2000).

SHOX2 represented a plausible candidate gene for JATD because of its expression pattern in tissues which are likely to be affected in JATD patients and because of its homology with *SHOX*, which is known to be mutated in a number of short stature syndromes characterized by skeletal abnormalities. However, no pathogenic mutations in the coding sequence of the gene were identified in the four JATD families screened. In the last stages of the project, DNA was obtained from an unaffected member of non-consanguineous Family 14 and, assuming linkage to 3q24-q26 in this family, *SHOX2* lies outside the interval identified by the recombination event which occurred in the unaffected child (see Table 4.23), thus excluding *SHOX2* as a candidate gene.

Twenty-eight known or predicted genes map to the 8 cM/12 Mb homozygous region common to Families 1, 5, 6, 8 and 14. While several of these genes can be ruled out on the grounds of inappropriate expression patterns and/or known roles in other human disorders, only eighteen of them were thought to represent possible candidate genes requiring further investigation. Many are genes encoding hypothetical proteins of unknown function and there are no obvious reasons for prioritising the screening of any of them.

5.4.2 Chromosome 8q24

5.4.2.1 WISP1

The *WISP1* gene, encoding the Wnt1-inducible signalling pathway protein 1 (WISP1/CCN4), belongs to the CCN family of secreted proteins and was thought to be a good candidate gene for JATD. *WISP1* maps to 8q24.1-q24.3 and lies within the overlapping homozygous interval in the two affected sibs from consanguineous Family 2, between markers D8S529 and D8S256 (see Figure 4.15).

The CCN family of secreted proteins (connective tissue growth factors) includes cysteine-rich 61 (CYR61/CCN1), connective tissue growth factor (CTGF/CCN2), nephroblastoma overexpressed (NOV/CCN3), Wnt1-inducible signalling pathway proteins-2 (WISP2/CCN5) and -3 (WISP3/CCN6). The CCN members are characterized by four conserved cysteine-rich domains: an insulin-like growth factor (IGF)-binding domain, a von Willebrand type C module, a thrombospondin-1 domain and a C-terminal domain containing a putative cysteine knot. Collectively, these proteins regulate fundamental biological processes, including cell adhesion, growth, differentiation, apoptosis, embryogenesis, vascular diseases and tumorigenesis and can interact with multiple signal transduction pathways as well as regulate other signalling molecules, including transforming growth factor β (TGF-β) and bone morphogenetic proteins (BMPs) during development. CCN members are known to be involved in skeletogenesis.

CYR61/CCN1 promotes collagen and cartilage production in limb buds and stimulates chondrogenesis, mitogenesis and adhesion in limb mesenchymal cells (Wong *et al.*, 1997). CTGF/CCN2 is expressed in zones containing hypertrophic chondrocytes or calcifying cartilage of long bones, ribs, the vertebral column and phalanges (Takigawa *et al.*, 2003). Increased levels of CTGF/CCN2 in skeletal cells are stimulated by TGF-β and BMP-2. CTGF/CCN2 is required for cell proliferation and matrix remodelling during chondrogenesis (Ivkovic *et al.*, 2003). *Ctgf* homozygous null mice die perinatally from respiratory failure as a consequence of impaired skeletal development.

Ctgf null mice suffer multiple skeletal dysmorphisms as a result of defective chondrocyte proliferation and decreased extracellular matrix components. Skeletal defects include sternum, rib cage and limb abnormalities (Ivkovic *et al.*, 2003).

WISP2/CCN5 is expressed at high levels in human bone tissue and plays an important role in regulating bone turnover (Kumar *et al.*, 1999). Mutations in the *WISP3/CCN6* gene are associated with progressive pseudorheumatoid dysplasia (PPAC; MIM 208230), an autosomal recessive skeletal disorder characterized by continued cartilage loss and bone destruction (Hurvitz *et al.*, 1999). Radiographic findings include osseous distension of the proximal and middle phalanges, enlargement of the femoral and tibial epiphyses and flattening of the vertebral bodies. *Wisp3* null mice have more mature vertebral endplates, a phenomenon that may be due to accelerated endochondral maturation (Perbal *et al.*, 2003).

WISP1/CCN4 was originally identified as a probable downstream effector of the Wnt signalling pathway in mouse mammary ephithelial cell lines transformed by Wnt-1 (Pennica *et al.*, 1998). In human tissues, *WISP1* expression was seen in a variety of adult organs including kidney, pancreas, lung, heart, ovary and spleen (Pennica *et al.*, 1998). High levels of expression have been demonstrated in fibroblast cells and colon tumors (Pennica *et al.*, 1998). Recently, the role of WISP1 during skeletal development has been elucidated and its expression during embryogenesis was shown in osteoblast and osteoblastic progenitor cells of the perichondral mesenchyme, making this gene an excellent functional candidate gene for JATD (French *et al.*, 2004).

WISP1 was screened for causative mutations in seventeen JATD families and although a number of SNPs were identified, no obvious mutations could be detected. The only changes found in the mother of the two affected sibs in Family 2 were two known SNPs. In one sporadic case (Family 31), a nucleotide substitution in exon 3 results in an amino acid change (Val184Ile). This is not likely to be significant, as it is a conservative change. Although mouse wisp1 has a valine at the same position, the same amino acid residue was not conserved in rat or chimp. In the proband from Family 24, a non-reported variant was found in the 5'UTR, 115 bp upstream of the start codon. This substitution is not listed in the NCBI SNP database. However, the surrounding region is not strongly conserved between human and mouse.

Following genotyping analysis of JATD families across the 8q24 interval, two more families were found to be consistent with linkage to this locus, Families 9 and 10. *WISP1* lies just outside the overlapping interval between the three families.

5.4.2.2 GLYLU

The novel gene *GLYLU* was the only gene in the interval of common homozygosity between Families 2, 9 and 10 and therefore, it was screened for mutations in one parent from each family. Although annotations about the protein (hypothetical protein LOC51059) expression and function were poor, *GLYLU* was considered an excellent positional candidate. It did not belong to any known protein family and had a predicted localization to the nucleus and a thiol protease site.

A number of polymorphisms were detected in the three individuals screened, but no convincing mutations such as a premature stop or a frameshifting insertion or deletion were identified in the coding region of *GLYLU*. The only change in the coding sequence of all three individuals was found in exon 8 and altered an amino acid (p.Arg362Ser). However, exon 8 was thought not to represent a true exon as it was predicted only by Ensembl and was not supported by evidence from any other source.

5.4.2.3 Further candidate genes on 8q24

As discussed in section 5.3.2, the diagnosis in the affected child from Family 9 was uncertain. Therefore, the attention in the search for candidate genes on 8q24 shifted to the region of common homozygosity in Families 2 and 10. Only three genes lay in this 3 Mb region between markers D8S256 and D8S274, *SIAT4A*, *ZNF406* and *KHDRBS3*.

ZNF406 was thought to represent the best candidate gene of the three and therefore it was the first one to be screened in Families 2 and 10. It contains multiple zinc finger motifs and appeared to encode a zinc finger transcription factor. It is expressed in a wide variety of adult tissues and tumours but no information was available about

whether it was expressed in cartilage/bone and during development. A number of human syndromes have been associated with defects in zinc finger proteins. Three related autosomal dominant disorders are caused by mutations in human *GLI3*, a zinc finger gene related to *Drosophila cubitus interruptus*. These are Greig cephalopolysyndactyly (GCPS; MIM 175700), Pallister-Hall syndrome (PHS; MIM 146510) and postaxial polysyndactyly type A1 (PAPA1; MIM 174200). GCPS is characterised by syndactyly, polydactyly, broad or bifid thumbs, hip dislocation, unusual skull shape, craniofacial dysmorphisms and developmental delay. PHS has a wide variety of features including virtually all of the findings in GCPS plus hypothalamic hamartoblastomas, hypopituarism, abnormalities of the palate, tongue, and jaw, limb shortening, postaxial polydactyly with nail dysplasia, syndactyly, fused ribs, imperforate anus, renal dysplasia and congenital heart defects. PAPA1 is characterised by an extra digit in the ulnar and/or fibular side of the upper and/or lower extremities. Hirschsprung disease (HSCR; MIM 235730) and Mowat-Wilson syndrome (MIM 235730) characterised by microcephaly, mental retardation, hypertelorism, cleft palate, heart defects and short stature are both caused by mutations in the zinc finger homeobox 1B gene (*ZFHX1B*). Mutations in *SALL4* and *SALL1*, two genes encoding zinc finger transcription factors, cause Okihiro syndrome (MIM 607323) and Townes-Brocks syndrome (TBS; MIM 107480), respectively. Patients affected with Okihiro syndrome have Duane retraction anomaly associated with radial ray abnormalities, whereas TBS is characterised by renal, anal, limb and ear anomalies.

ZNF406 was therefore sequenced in unaffected parents from Families 2 and 10 but no pathogenic mutations were identified. The only significant change in Family 2 was a p.Gly105Arg identified in exon 3. This alteration is not listed in the NCBI SNP database and it is not a conservative change. Moreover, the Gly residue is conserved in the mouse protein. The variant segregates in the family consistently with a possible pathogenic role, although the sequence of one human mRNA IMAGE clone has the same nucleotide alteration. Furthermore, the substitution does not occur in a region of the gene of obvious functional significance and is not conserved in fugu, therefore it is likely to represent a rare polymorphism.

Due to time constraints, the two other genes in the region, *SIAT4A* and *KHDRBS3*, were not screened. Both represent good candidate genes and need to be screened in the future.

SIAT4A (sialyltransferase 4A) is a glycosyltransferase which catalyses the transfer of sialic acid to terminal carbohydrate groups of cell-surface glycolipids and glycoproteins and is therefore important for cell-cell interactions. It is widely expressed in adult tissues, cancer and fetal tissues. It is expressed most strongly in skeletal muscle, but also in tissues relevant to JATD, such as liver, kidney, pancreas, retina and cartilage. There are two known transcripts produced by alternative splicing containing either nine or ten exons.

KHDRBS3, alternatively named *T-STAR* (testis-signal transduction and activation of RNA), is a Sam68-like phosphotyrosine nuclear protein involved in RNA processing and signal transduction specifically involved in regulating alternative splicing in response to extra-cellular signals. It contains an RNA binding domain, six SH3-binding sites, and a tyrosine-rich C-terminus. It is expressed in testis, skeletal muscle, brain and heart and it is represented in databases by cDNAs from a wide variety of adult tissues and tumours including liver, kidney and retina.

5.4.3 Chromosome 14q24-q32: *TTC7B*

It has been suggested that genes encoding intraflagellar transport (IFT) proteins might be promising candidate disease genes for human disorders such as JATD (Pazour and Rosenbaum, 2002; Rosenbaum and Witman, 2002). IFT is a microtubule-dependent motility process necessary for cilia biogenesis and maintenance (Rosenbaum and Witman, 2002; Scholey, 2003; Qin *et al.*, 2004). Defects in IFT affect the cilia in several organs leading to a pleiotrophic phenotype affecting several organs systems like the kidney and the eye in humans (Rosenbaum and Witman, 2002). In recent years, it has been demonstrated that defects in cilia are associated with several human disorders, including nephronophthisis (NPHP; MIM 608002), hydrocephalus, polycystic kidney disease (PKD; MIM 173900), Senior-Loken syndrome (MIM 266900) and Bardet-

Biedl syndrome (BBS; MIM 209900) (Watnick and Germino, 2003). Some of the phenotypic features exhibited by JATD patients, including skeletal defects associated with renal cysts/malformations and retinal degeneration, resemble human diseases which have been associated with abnormal cilia function such as NPHP, PKD and BBS. These features are similar to the phenotype of the KO mouse *Tg737* (Pazour and Rosenbaum, 2002; Watnick and Germino, 2003). In the mutant mouse *Tg737* the primary defect is the inability to assemble cilia. The *Tg737* gene encodes a protein (Polaris) homologous to the IFT88 subunit of the *Chlamydomonas* IFT particle, required for flagellar assembly in *Chlamydomonas* (Pazour *et al.*, 2000). The *Tg737* mouse mutant develops PKD, retinal degeneration and laterality defects (*situs inversus*) as well as hepatic fibrosis, liver cysts, polydactyly and hydrocephaly (Pazour *et al.*, 2000).

In JATD patients, Ozcay and colleagues described two different patterns of renal disease. Patients who died in the neonatal period or during infancy exhibit various types of renal dysplasia with cortical or diffuse cystic changes, while those who survive beyond childhood develop progressive renal disease with diffuse tubulointerstitial fibrous changes and infiltration, tubular atrophy and periglomerular fibrosis or glomerular sclerosis (Ozcay *et al.*, 2001). Several reports have claimed that these findings are indistinguishable from those associated with juvenile NPHP (Kozlowski and Masel, 1976; Shah, 1980; Donaldson *et al.*, 1985).

A feature of JATD is retinal degeneration resembling Leber congenital amaurosis and tapeto retinal degeneration like those described in Senior-Loken syndrome, BBS and NPHP. Hypertrophic retinopathy has also been described (Brueton *et al.*, 1990; Ozcay *et al.*, 2001).

Laterality defects caused by a defective cilia assembly in the embryonic node cells are also often associated with cilia dysfunction diseases such as Kartagener's syndrome (MIM 244400; characterized by *situs inversus* and respiratory disease) and BBS (Ansley *et al.*, 2003). *Situs inversus* has been described in two JATD cases (Brueton *et al.*, 1990; Majewski *et al.*, 1996). Hepatic fibrosis, liver cysts and polydactyly are often associated with JATD, whereas mild congenital hydrocephalus has been described in two sibs with JATD (Singh *et al.*, 1988).

TTC7B encoding the tetratricopeptide repeats domain 7B protein lies on 14q24-q32 and was thought to be a good candidate gene for JATD. Although at present there is no evidence to support a possible role of TTC7B in IFT, the structure of the predicted protein suggested a possible function of TTC7B in ciliary function, IFT, ciliogenesis or in the cell cycle, similar to that of BBS8 and BBS4. TTC7B contains five TPRs and a domain similar to the prokaryotic *pilF* domain involved in pilus assembly and intracellular trafficking. BBS8 also contains eight TPRs towards the C-terminus and has a *pilF* domain. It localizes to the centrosomes and basal bodies in cells interacting with PCM-1, a protein probably involved in ciliogenesis (Ansley *et al.*, 2003). BBS8 localizes specifically to ciliated structures, such as the connecting cilium of the retina and columnar epithelial cells in the lung. In one family, null *BBS8* mutation leads to BBS with *situs inversus* (Ansley *et al.*, 2003). BBS4 contains ten TPRs and is a pericentriolar protein necessary in recruiting cargo to centriolar satellites by acting as a dynein-associated adaptor protein (Kim *et al.*, 2004).

No mutations were identified in the coding sequence of *TTC7B*, although four exons have not been yet screened for mutations.

5.5 Conclusions

In conclusion, the investigation into the molecular basis of JATD carried out in this study led to the identification of three putative JATD loci on chromosome 3q24-q26, 8q24 and 14q24-q32 through analysis of a number of consanguineous and non-consanguineous JATD families. Locus heterogeneity has therefore been demonstrated. Two families described in this project and several families ascertained by our collaborators do not appear to be consistent with linkage to any of the known loci, suggesting that there might be at least one additional and as yet unidentified locus.

Some families exhibited homozygosity across two JATD putative loci. The most likely explanation is that these families display spurious linkage across one of the two regions, thus the homozygosity in that region is IBS and not IBD. Alternatively, it might suggest a more complex pattern of inheritance such as multilocus inheritance. It

is in fact possible that mutations at two or more loci are necessary for pathogenesis of JATD. The multiallelic model of inheritance is a widespread phenomenon and examples have been reported in a number of species including humans (Badano and Katsanis, 2002; Ming and Muenke, 2002). In the case of Bardet-Biedl syndrome, for example, the model of disease transmission postulated has been called triallelic inheritance, as three mutated alleles from two loci are required for penetrance of the BBS phenotype in some families (Katsanis *et al.*, 2001; Katsanis *et al.*, 2002; Beales *et al.*, 2003). As demonstrated in this study, as in BBS, JATD is a rare genetically heterogeneous disorder characterized by multiple clinical features with variable expressivity. At present, it is difficult to explain how the disruption of a single protein would lead to such a complex phenotype as in JATD that includes both skeletal dysplasias (e.g. small rib cage, short stature, rhizomelic limb shortening and polydactyly) and progressive renal, pancreatic, hepatic and retinal defects. At present the hypothesis of a complex transmission model for JATD cannot be substantiated and only the identification of the underlying genes and the pathogenic mutations in them will help to elucidate this.

Mutation screening of the most promising candidate genes lying in the three putative loci has been carried out in this study. Although all the genes screened appeared to represent excellent candidates, no pathogenic mutation such as a premature stop, an insertion, a deletion, or a splice-site mutation, could be identified in any of them. In general, it is possible that mutations in the coding sequence of the genes screened have been missed, although this is very unlikely as the analysis has been thoroughly carried out. It is known that DNA variants could lie in regulatory regions altering the consensus of transcription factor binding sites or promoter elements. As previously described (section 3.3.4) an example is given by point mutations identified in a *Shh* regulatory element located within intron 5 of the *Lmbr1* gene, 1 Mb upstream of *Shh* itself, in human preaxial polydactyly (Lettice *et al.*, 2003). Alternatively, variants in the untranslated regions of mRNA may alter mRNA stability. Mutations within introns, regulatory elements, 5' and 3'UTRs, or that create or disrupt cryptic splice sites would have been missed by the methodology used in this study.

Another possibility, especially in novel genes, is that there are other spliced variants that have been missed and not analysed. Moreover, not all the families were screened for mutations in all the genes and therefore it is possible that mutations have been missed. For example, no mutations in the coding sequence of *SHOX2* were identified in the four families consistent with linkage to 3q24-q26, although the gene was not screened in the sporadic cases. On 8q24, *GLYLU* and *ZNF406* were sequenced only in the families consistent with linkage to this locus and not in the sporadic cases.

Alternatively, some of the polymorphisms identified in coding and non-coding regions of the genes analysed could be significant. A number of functional polymorphisms in gene regulatory sequences have in fact been described. For example, polymorphisms that appear to be innocuous on cursory examination can have functional consequences, such as synonymous coding changes that occur in exon splicing enhancer (ESE) regions, as described for the breast cancer coding gene *BRCA1* (Liu *et al.*, 2001).

Finally, it is possible that the genes underlying JATD are between the ones that have not yet been screened.

5.6 Future work

A number of JATD families described in this study need to be investigated further across the five JATD putative loci. More informative markers should be genotyped in Families 4 and 16 on chromosome 3q24-q26. Family 4 in particular might be useful to narrow down the region if linkage can be supported. More typing is necessary in Families 6 and 16 on chromosome 12 as the results obtained so far have been inconclusive. Family 5 needs to be analysed across the chromosome 8 and 12 regions. To support the hypothesis of linkage to either the 3q24-q26 or the 15q13 locus in Family 7, it would be necessary to extract DNA from the paraffin embedded tissues from the affected child. Genotyping of the child might support or exclude linkage to one or both loci. The region identified on 14q24-q32 in Family 4 requires further study and, if it is proven to be a real region homozygous by descent, has to be investigated in all the other families. Furthermore, linkage to the previously reported candidate loci

such as the *EVC/EVC2* locus, 12p11.21-p12.2 and the regions syntenic to the *srt* locus was not excluded in all families.

More JATD families should be ascertained and genotyped across the five putative JATD loci to validate and refine the candidate regions as tightly as possible, thus providing the basis for further candidate gene analysis.

An approach that could be adopted is the search for allelic homozygosity in non-consanguineous families within the candidate regions as described by Morgan and colleagues (Morgan *et al.*, 2003). Some of the non-consanguineous families might in fact show homozygosity within the critical interval due to distant consanguinity. This approach could be useful to narrow down the putative loci as the homozygous interval in such families is expected to be small, although it could be time consuming if large numbers of families are to be analysed. A large number of closely spaced polymorphic markers or SNPs would be required for this kind of study.

A possible approach to validate candidate linked regions or to identify new loci in consanguineous families would be to perform homozygosity mapping using new technologies such as the Affimetrix GeneChip (Kennedy *et al.*, 2003). This high throughput technique allows genotyping of 10,000 or 100,000 SNPs simultaneously per single target DNA using allele specific hybridisation. The advantage of using SNPs for genome screening instead of traditional microsatellite markers is that SNPs are abundant, widespread throughout the genome and represent the most frequent form of DNA sequence variation found in the human genome with an average spacing of one SNP every 210 Kb according to the November 2002 NCBI genome assembly (Build 33). The average spacing of microsatellite markers is about one every 5.6 Mb.

Mutation screening of further candidate genes on chromosomes 3, 8 and 14 has to be performed. Recent scrutiny of the genomic sequence on 8q24 in the region of triple overlap between Families 2, 9 and 10 using Ensembl, UCSC genome browser and NCBI Map View (Build 35, May 2004 freeze) led to the identification of two novel genes (FLJ45872 and LOC442397). These genes were not annotated at the time of this study and need to be sequenced in the affected families. *SIAT4A* and *KHDRBS3* lying in the region of common homozygosity in Families 2 and 10 on 8q24 need to be

screened. The region on 3q24-q26 is still quite large with a possible eighteen candidate genes. Therefore genes for mutation screening need to be prioritised based on their biological rationale in the target phenotype.

Genes can be prioritised based on expression in the disease tissues, a known or putative role in the disease pathway and gene knock-out models. The candidate genes should be expressed in the right place at the right time for the disease phenotype. Information on gene expression can be obtained by searching databases such as dbEST. RT-PCR, Northern blotting, Serial Analysis of Gene Expression (SAGE) and DNA microarrays are all techniques available to assess expression of candidate genes. The genes underlying JATD are most likely to be expressed in cartilage, therefore a possible approach to selecting plausible candidate genes in each region would be to investigate their expression in human primary chondrocytes by using an Affymetrix microarray system. Using a Human Genome U133 GeneChip set, a comprehensive profile of gene expression in this cell type could be obtained. This chip allows over 33,000 proven and potential transcripts to be quantitated relative to a set of 100 housekeeping genes with the lower limit of detection roughly corresponding to three transcripts per cell. For this purpose, RNA from human fetal primary chondrocytes has to be obtained.

The candidate genes and/or their proteins should have an appropriate function for the disease. Sometimes function of novel genes is difficult to establish as there are very few annotations in the databases. Clues to prioritise candidate genes may surface by examining the functional relationship to genes known to be involved in similar diseases. In the case of JATD for example, studies on the function of *EVC* and *EVC2* might give useful hints about the possible function of the JATD underlying genes. Genes encoding IFT proteins and proteins involved in ciliogenesis are also promising candidate disease genes for JATD. It is therefore necessary to search the putative JATD loci for genes encoding proteins containing domains known to be involved in these processes such as TPRs and pilF domains. Moreover, the identification of the genes underlying one or more short rib-polydactyly syndromes (SRPS) could also be valuable to suggest other candidate genes to investigate. Similarities in the phenotype of SRPS and JATD are in fact striking.

A mouse mutant phenotypically similar to JATD could also be valuable in suggesting candidate genes to screen. Finally, it is possible that proteins encoded by the JATD genes either belong to the same family of proteins or are part of a common developmental pathway or interact with one another in multi-aggregates. In these cases, the identification of at least one gene in one of the JATD loci may direct the search for genes in the other putative loci.

The identification of the JATD genes will provide an important new insight into retinal, limb and kidney development.

References

Afzal, A. R., and Jeffery, S. (2003). One gene, two phenotypes: *ROR2* mutations in autosomal recessive Robinow syndrome and autosomal dominant brachydactyly type B. Hum Mutat *22*, 1-11.

Afzal, A. R., Rajab, A., Fenske, C., Crosby, A., Lahiri, N., Ternes-Pereira, E., Murday, V. A., Houlston, R., Patton, M. A., and Jeffery, S. (2000a). Linkage of recessive Robinow syndrome to a 4 cM interval on chromosome 9q22. Hum Genet *106*, 351-354.

Afzal, A. R., Rajab, A., Fenske, C. D., Oldridge, M., Elanko, N., Ternes-Pereira, E., Tuysuz, B., Murday, V. A., Patton, M. A., Wilkie, A. O., and Jeffery, S. (2000b). Recessive Robinow syndrome, allelic to dominant brachydactyly type B, is caused by mutation of *ROR2*. Nat Genet *25*, 419-422.

Ansley, S. J., Badano, J. L., Blacque, O. E., Hill, J., Hoskins, B. E., Leitch, C. C., Kim, J. C., Ross, A. J., Eichers, E. R., Teslovich, T. M., *et al.* (2003). Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. Nature *425*, 628-633.

Armstrong, J. F., Pritchard-Jones, K., Bickmore, W. A., Hastie, N. D., and Bard, J. B. (1993). The expression of the Wilms' tumour gene, *WT1*, in the developing mammalian embryo. Mech Dev *40*, 85-97.

Bacchelli, C., Goodman, F. R., Scambler, P. J., and Winter, R. M. (2001). Cenani-Lenz syndrome with renal hypoplasia is not linked to *FORMIN* or *GREMLIN*. Clinical Genetics *59*, 203-205.

Badano, J. L., Ansley, S. J., Leitch, C. C., Lewis, R. A., Lupski, J. R., and Katsanis, N. (2003). Identification of a novel Bardet-Biedl syndrome protein, BBS7, that shares structural features with BBS1 and BBS2. Am J Hum Genet *72*, 650-658.

Badano, J. L., and Katsanis, N. (2002). Beyond Mendel: an evolving view of human genetic disease transmission. Nat Rev Genet *3*, 779-789.

Bahring, S., Nagai, T., Toka, H. R., Nitz, I., Toka, O., Aydin, A., Muhl, A., Wienker, T. F., Schuster, H., and Luft, F. C. (1997). Deletion at 12p in a Japanese child with brachydactyly overlaps the assigned locus of brachydactyly with hypertension in a Turkish family. Am J Hum Genet *60*, 732-735.

Ballabio, A. (1993). The rise and fall of positional cloning? Nat Genet *3*, 277-279.

Bamshad, M., Lin, R. C., Law, D. J., Watkins, W. C., Krakowiak, P. A., Moore, M. E., Franceschini, P., Lala, R., Holmes, L. B., Gebuhr, T. C., *et al.* (1997). Mutations in human *TBX3* alter limb, apocrine and genital development in ulnar-mammary syndrome. Nat Genet *16*, 311-315.

Basson, C. T., Bachinsky, D. R., Lin, R. C., Levi, T., Elkins, J. A., Soults, J., Grayzel, D., Kroumpouzou, E., Traill, T. A., Leblanc-Straceski, J., *et al.* (1997). Mutations in human *TBX5* [corrected] cause limb and cardiac malformation in Holt-Oram syndrome. Nat Genet *15*, 30-35.

Beales, P. L., Badano, J. L., Ross, A. J., Ansley, S. J., Hoskins, B. E., Kirsten, B., Mein, C. A., Froguel, P., Scambler, P. J., Lewis, R. A., *et al.* (2003). Genetic interaction of *BBS1* mutations with alleles at other *BBS* loci can result in non-Mendelian Bardet-Biedl syndrome. Am J Hum Genet *72*, 1187-1199.

Ben Hamida, C., Doerflinger, N., Belal, S., Linder, C., Reutenauer, L., Dib, C., Gyapay, G., Vignal, A., Le Paslier, D., Cohen, D., and et al. (1993). Localization of Friedreich ataxia phenotype with selective vitamin E deficiency to chromosome 8q by homozygosity mapping. Nat Genet *5*, 195-200.

Blacque, O. E., Reardon, M. J., Li, C., McCarthy, J., Mahjoub, M. R., Ansley, S. J., Badano, J. L., Mah, A. K., Beales, P. L., Davidson, W. S., *et al.* (2004). Loss of C. elegans BBS-7 and BBS-8 protein function results in cilia defects and compromised intraflagellar transport. Genes Dev *18*, 1630-1642.

Blaschke, R. J., Monaghan, A. P., Schiller, S., Schechinger, B., Rao, E., Padilla-Nash, H., Ried, T., and Rappold, G. A. (1998). *SHOT*, a *SHOX*-related homeobox gene, is implicated in craniofacial, brain, heart, and limb development. Proc Natl Acad Sci U S A *95*, 2406-2411.

Bosse, K., Betz, R. C., Lee, Y. A., Wienker, T. F., Reis, A., Kleen, H., Propping, P., Cichon, S., and Nothen, M. M. (2000). Localization of a gene for syndactyly type 1 to chromosome 2q34-q36. Am J Hum Genet *67*, 492-497.

Bouchard, M., Souabni, A., Mandler, M., Neubuser, A., and Busslinger, M. (2002). Nephric lineage specification by Pax2 and Pax8. Genes Dev *16*, 2958-2970.

Boulet, A.M., Moon, A.M., Arenkiel, B.R., and Capecchi, M.R. (2004). The roles of Fgf4 and Fgf8 in limb bud initiation and outgrowth. Dev Biol *15*;273, 361-372

Brueton, L. A., Dillon, M. J., and Winter, R. M. (1990). Ellis-van creveld syndrome, Jeune syndrome, and renal-hepatic- pancreatic dysplasia: separate entities or disease spectrum? J Med Genet *27*, 252-255.

Bulman, M. P., Kusumi, K., Frayling, T. M., McKeown, C., Garrett, C., Lander, E. S., Krumlauf, R., Hattersley, A. T., Ellard, S., and Turnpenny, P. D. (2000). Mutations in the human delta homologue, *DLL3*, cause axial skeletal defects in spondylocostal dysostosis. Nat Genet *24*, 438-441.

Capilupi, B., Olappi, G., Cornaglia, A. M., and Novati, G. P. (1996). [Asphyxiating thoracic dysplasia or Jeune's syndrome. Description of 2 mild familial cases]. Pediatr Med Chir *18*, 529-532.

Carmi, R., Rokhlina, T., Kwitek-Black, A. E., Elbedour, K., Nishimura, D., Stone, E. M., and Sheffield, V. C. (1995). Use of a DNA pooling strategy to identify a human obesity syndrome locus on chromosome 15. Hum Mol Genet *4*, 9-13.

Casteels, I., Demandt, E., and Legius, E. (2000). Visual loss as the presenting sign of Jeune syndrome. Europ J Paediatr Neurol *4*, 243-247.

Cenani, A., and Lenz, W. (1967). [Total syndactylia and total radioulnar synostosis in 2 brothers. A contribution on the genetics of syndactylia]. Z Kinderheilkd *101*, 181-190.

Chan, D. C., Bedford, M. T., and Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. Embo J *15*, 1045-1054.

Chan, D. C., and Leder, P. (1996). Genetic evidence that formins function within the nucleus. J Biol Chem *271*, 23472-23477.

Chan, D. C., Wynshaw-Boris, A., and Leder, P. (1995). Formin isoforms are differentially expressed in the mouse embryo and are required for normal expression of *fgf-4* and shh in the limb bud. Development *121*, 3151-3162.

Clement-Jones, M., Schiller, S., Rao, E., Blaschke, R. J., Zuniga, A., Zeller, R., Robson, S. C., Binder, G., Glass, I., Strachan, T., *et al.* (2000). The short stature homeobox gene *SHOX* is involved in skeletal abnormalities in Turner syndrome. Hum Mol Genet *9*, 695-702.

Collins, F. S. (1992). Positional cloning: let's not call it reverse anymore. Nat Genet *1*, 3-6.

Collins, F. S. (1995). Positional cloning moves from perditional to traditional. Nat Genet *9*, 347-350.

De Smet, L., De Beer, P., and Fryns, J. P. (1996). Cenani-Lenz syndrome in father and daughter. Genet Couns *7*, 153-157.

De Smet, L., Winnepenninckx, B., Fryns, J. P., and Fabry, G. (1992). Cenani-Lenz type of syndactyly: a complex type of syndactyly with multiple synostoses. Genet Couns *3*, 145-147.

Debeer, P., Schoenmakers, E. F., De Smet, L., Van de Ven, W. J., and Fryns, J. P. (1998a). Co-segregation of an apparently balanced reciprocal t(12;22)(p11.2;q13.3) with a complex type of 3/3'/4 synpolydactyly associated with metacarpal, metatarsal and tarsal synostoses in three family members. Clin Dysmorphol *7*, 225-228.

Debeer, P., Schoenmakers, E. F., Thoelen, R., Fryns, J. P., and Van de Ven, W. J. (1998b). Physical mapping of the t(12;22) translocation breakpoints in a family with a complex type of 3/3'/4 synpolydactyly. Cytogenet Cell Genet *81*, 229-234.

Debeer, P., Schoenmakers, E. F., Thoelen, R., Holvoet, M., Kuittinen, T., Fabry, G., Fryns, J. P., Goodman, F. R., and Van de Ven, W. J. (2000). Physical map of a 1.5 mb region on 12p11.2 harbouring a synpolydactyly associated chromosomal breakpoint. Eur J Hum Genet *8*, 561-570.

Debeer, P., Schoenmakers, E. F., Twal, W. O., Argraves, W. S., De Smet, L., Fryns, J. P., and Van De Ven, W. J. (2002). The fibulin-1 gene (*FBLN1*) is disrupted in a t(12;22) associated with a complex type of synpolydactyly. J Med Genet *39*, 98-104.

den Dunnen, J. T., and Antonarakis, S. E. (2000). Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat *15*, 7-12.

Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., *et al.* (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature *380*, 152-154.

DiLella, A. G., Kwok, S. C., Ledley, F. D., Marvit, J., and Woo, S. L. (1986). Molecular structure and polymorphic map of the human phenylalanine hydroxylase gene. Biochemistry *25*, 743-749.

Dodinval, P. (1979). Oligodactyly and multiple synostoses of the extremities: two cases in sibs. A variant of Cenani-Lenz syndactyly. Hum Genet *48*, 183-189.

Dolan, V., Murphy, M., Alarcon, P., Brady, H. R., and Hensey, C. (2003). Gremlin - a putative pathogenic player in progressive renal disease. Expert Opin Ther Targets *7*, 523-526.

Donaldson, M. D., Warner, A. A., Trompeter, R. S., Haycock, G. B., and Chantler, C. (1985). Familial juvenile nephronophthisis, Jeune's syndrome, and associated disorders. Arch Dis Child *60*, 426-434.

Drohm, D., Lenz, W., and Yang, T. S. (1976). [Total syndactylism with mesomelic shortening of the arm, radioulnar and metacarpal synostoses and disorganization of the phalanges ("cenani syndactylism") (author's transl)]. Klin Padiatr *188*, 359-365.

Dudley, A. T., Ros, M. A., and Tabin, C. J. (2002). A re-examination of proximodistal patterning during vertebrate limb development. Nature *418*, 539-544.

Eden, E. R., Naoumova, R. P., Burden, J. J., McCarthy, M. I., and Soutar, A. K. (2001). Use of homozygosity mapping to identify a region on chromosome 1 bearing a defective gene that causes autosomal recessive homozygous hypercholesterolemia in two unrelated families. Am J Hum Genet *68*, 653-660.

Elcioglu, N., Atasu, M., and Cenani, A. (1997). Dermatoglyphics in patients with Cenani-Lenz type syndactyly: studies in a new case. Am J Med Genet *70*, 341-345.

Esquela, A. F., and Lee, S. J. (2003). Regulation of metanephric kidney development by growth/differentiation factor 11. Dev Biol *257*, 356-370.

Evangelista, M., Pruyne, D., Amberg, D. C., Boone, C., and Bretscher, A. (2002). Formins direct Arp2/3-independent actin filament assembly to polarize cell growth in yeast. Nat Cell Biol *4*, 32-41.

Evangelista, M., Zigmond, S., and Boone, C. (2003). Formins: signaling effectors for assembly and polarization of actin filaments. J Cell Sci *116*, 2603-2611.

Finegold, M. J., Katzew, H., Genieser, N. B., and Becker, M. H. (1971). Lung structure in thoracic dystrophy. Am J Dis Child *122*, 153-159.

Franceschini, P., Guala, A., Vardeu, M. P., Signorile, F., Franceschini, D., and Bolgiani, M. P. (1995). Short rib-dysplasia group (with/without polydactyly): report of a patient suggesting the existence of a continuous spectrum. Am J Med Genet *59*, 359-364.

French, D. M., Kaul, R. J., D'Souza, A. L., Crowley, C. W., Bao, M., Frantz, G. D., Filvaroff, E. H., and Desnoyers, L. (2004). WISP-1 Is an Osteoblastic Regulator Expressed During Skeletal Development and Fracture Repair. Am J Pathol *165*, 855-867.

Fromenty, B., Demeilliers, C., Mansouri, A., and Pessayre, D. (2000). Escherichia coli exonuclease III enhances long PCR amplification of damaged DNA templates. Nucleic Acids Res *28*, E50.

Galdzicka, M., Patnala, S., Hirshman, M. G., Cai, J. F., Nitowsky, H., Egeland, J. A., and Ginns, E. I. (2002). A new gene, *EVC2*, is mutated in Ellis-van Creveld syndrome. Mol Genet Metab *77*, 291-295.

Genin, E., Todorov, A. A., and Clerget-Darpoux, F. (1998). Optimization of genome search strategies for homozygosity mapping: influence of marker spacing on power and threshold criteria for identification of candidate regions. Ann Hum Genet *62*, 419-429.

Georgiou-Theodoropoulos, M., Agapitos, M., Theodoropoulos, P., and Koutselinis, A. (1988). Jeune syndrome associated with pancreatic fibrosis. Pediatr Pathol *8*, 541-544.

Giorgi, P. L., Gabrielli, O., Bonifazi, V., Catassi, C., and Coppa, G. V. (1990). Mild form of Jeune syndrome in two sisters. Am J Med Genet *35*, 280-282.

Gitschier, J., Wood, W. I., Goralka, T. M., Wion, K. L., Chen, E. Y., Eaton, D. H., Vehar, G. A., Capon, D. J., and Lawn, R. M. (1984). Characterization of the human factor VIII gene. Nature *312*, 326-330.

Grieshammer, U., Le, M., Plump, A. S., Wang, F., Tessier-Lavigne, M., and Martin, G. R. (2004). SLIT2-mediated ROBO2 signaling restricts kidney induction to a single site. Dev Cell *6*, 709-717.

Haider, N. B., Carmi, R., Shalev, H., Sheffield, V. C., and Landau, D. (1998). A Bedouin kindred with infantile nephronophthisis demonstrates linkage to chromosome 9 by homozygosity mapping. Am J Hum Genet *63*, 1404-1410.

Herron, B. J., Lu, W., Rao, C., Liu, S., Peters, H., Bronson, R. T., Justice, M. J., McDonald, J. D., and Beier, D. R. (2002). Efficient generation and mapping of recessive developmental mutations using ENU mutagenesis. Nat Genet *30*, 185-189.

Ho, N. C., Francomano, C. A., and van Allen, M. (2000). Jeune asphyxiating thoracic dystrophy and short-rib polydactyly type III (Verma-Naumoff) are variants of the same disorder. Am J Med Genet *90*, 310-314.

Hopper, M. S., Boultbee, J. E., and Watson, A. R. (1979). Polyhydramnios associated with congenital pancreatic cysts and asphyxiating thoracic dysplasia. A case report. S Afr Med J *56*, 32-33.

Howard, T. D., Guttmacher, A. E., McKinnon, W., Sharma, M., McKusick, V. A., and Jabs, E. W. (1997). Autosomal dominant postaxial polydactyly, nail dystrophy, and dental abnormalities map to chromosome 4p16, in the region containing the Ellis-van Creveld syndrome locus. Am J Hum Genet *61*, 1405-1412.

Hsu, D. R., Economides, A. N., Wang, X., Eimon, P. M., and Harland, R. M. (1998). The Xenopus dorsalizing factor Gremlin identifies a novel family of secreted proteins that antagonize BMP activities. Mol Cell *1*, 673-683.

Hudgins L, R. S., Treem W, Hyams J (1990). Early cirrhosis in survivors with Jeune thoracic dystrophy. Am J Hum Genet *47*, A61.

Hurvitz, J. R., Suwairi, W. M., Van Hul, W., El-Shanti, H., Superti-Furga, A., Roudier, J., Holderbaum, D., Pauli, R. M., Herd, J. K., Van Hul, E. V., *et al.* (1999). Mutations in the CCN gene family member *WISP3* cause progressive pseudorheumatoid dysplasia. Nat Genet *23*, 94-98.

Ivkovic, S., Yoon, B. S., Popoff, S. N., Safadi, F. F., Libuda, D. E., Stephenson, R. C., Daluiski, A., and Lyons, K. M. (2003). Connective tissue growth factor coordinates chondrogenesis and angiogenesis during skeletal development. Development *130*, 2779-2791.

Jackson-Grusby, L., Kuo, A., and Leder, P. (1992). A variant limb deformity transcript expressed in the embryonic mouse limb defines a novel formin. Genes Dev *6*, 29-37.

Jankauskiene, A., and Bernatoniene, J. (2000). Clinical quiz. Jeune syndrome. Pediatr Nephrol *14*, 1054-1056.

Jeune M, B. C., Carron R (1955). Dystrophie thoracique asphyxiante de caracter familial. Archives Francaises de Pediatrie *12*, 886-891.

Jeune M, C. R., Beraud C, Loaec Y (1954). Polychondrodystrophie avec blocage thoracique d'evolution fatale. Pediatrie *9*, 390-392.

Kaddoura, I. L., Obeid, M. Y., Mroueh, S. M., and Nasser, A. A. (2001). Dynamic thoracoplasty for asphyxiating thoracic dystrophy. Ann Thorac Surg *72*, 1755-1758.

Kajantie, E., Andersson, S., and Kaitila, I. (2001). Familial asphyxiating thoracic dysplasia: clinical variability and impact of improved neonatal intensive care. J Pediatr *139*, 130-133.

Kalatzis, V., and Petit, C. (1998). The fundamental and medical impacts of recent progress in research on hereditary hearing loss. Hum Mol Genet *7*, 1589-1597.

Katoh, M. (2004a). Identification and characterization of human *DIAPH3* gene in silico. Int J Mol Med *13*, 473-478.

Katoh, M. (2004b). Identification and characterization of human *FHOD3* gene in silico. Int J Mol Med *13*, 615-620.

Katsanis, N., Ansley, S. J., Badano, J. L., Eichers, E. R., Lewis, R. A., Hoskins, B. E., Scambler, P. J., Davidson, W. S., Beales, P. L., and Lupski, J. R. (2001). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. Science *293*, 2256-2259.

Katsanis, N., Eichers, E. R., Ansley, S. J., Lewis, R. A., Kayserili, H., Hoskins, B. E., Scambler, P. J., Beales, P. L., and Lupski, J. R. (2002). *BBS4* is a minor contributor to Bardet-Biedl syndrome and may also participate in triallelic inheritance. Am J Hum Genet *71*, 22-29.

Kennedy, G. C., Matsuzaki, H., Dong, S., Liu, W. M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J., *et al.* (2003). Large-scale genotyping of complex DNA. Nat Biotechnol *21*, 1233-1237.

Kerem, B., Rommens, J. M., Buchanan, J. A., Markiewicz, D., Cox, T. K., Chakravarti, A., Buchwald, M., and Tsui, L. C. (1989). Identification of the cystic fibrosis gene: genetic analysis. Science *245*, 1073-1080.

Khokha, M. K., Hsu, D., Brunet, L. J., Dionne, M. S., and Harland, R. M. (2003). Gremlin is the BMP antagonist required for maintenance of Shh and Fgf signals during limb patterning. Nat Genet *34*, 303-307.

Kim, J. C., Badano, J. L., Sibold, S., Esmail, M. A., Hill, J., Hoskins, B. E., Leitch, C. C., Venner, K., Ansley, S. J., Ross, A. J., *et al.* (2004). The Bardet-Biedl protein BBS4 targets cargo to the pericentriolar region and is required for microtubule anchoring and cell cycle progression. Nat Genet *36*, 462-470.

Kobielak, A., Pasolli, H. A., and Fuchs, E. (2004). Mammalian formin-1 participates in adherens junctions and polymerization of linear actin cables. Nat Cell Biol *6*, 21-30.

Koka, S., Neudauer, C. L., Li, X., Lewis, R. E., McCarthy, J. B., and Westendorf, J. J. (2003). The formin-homology-domain-containing protein FHOD1 enhances cell migration. J Cell Sci *116*, 1745-1755.

Kozlowski, K., and Masel, J. (1976). Asphyxiating thoracic dystrophy without respiratory disease: report of two cases of the latent form. Pediatr Radiol *5*, 30-33.

Krakow, D., Salazar, D., Wilcox, W. R., Rimoin, D. L., and Cohn, D. H. (2000). Exclusion of the Ellis-van Creveld region on chromosome 4p16 in some families with asphyxiating thoracic dystrophy and short-rib polydactyly syndromes. Eur J Hum Genet *8*, 645-648.

Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. Cell *74*, 679-691.

Kulaga, H. M., Leitch, C. C., Eichers, E. R., Badano, J. L., Lesemann, A., Hoskins, B. E., Lupski, J. R., Beales, P. L., Reed, R. R., and Katsanis, N. (2004). Loss of BBS proteins causes anosmia in humans and defects in olfactory cilia structure and function in the mouse. Nat Genet *36*, 994-998.

Kumar, S., Hand, A. T., Connor, J. R., Dodds, R. A., Ryan, P. J., Trill, J. J., Fisher, S. M., Nuttall, M. E., Lipshutz, D. B., Zou, C., *et al.* (1999). Identification and cloning of a connective tissue growth factor-like cDNA from human osteoblasts encoding a novel regulator of osteoblast functions. J Biol Chem *274*, 17123-17131.

Kusumi, K., Sun, E. S., Kerrebrock, A. W., Bronson, R. T., Chi, D. C., Bulotsky, M. S., Spencer, J. B., Birren, B. W., Frankel, W. N., and Lander, E. S. (1998). The mouse pudgy mutation disrupts *Delta* homologue *Dll3* and initiation of early somite boundaries. Nat Genet *19*, 274-278.

Kwitek-Black, A. E., Carmi, R., Duyk, G. M., Buetow, K. H., Elbedour, K., Parvari, R., Yandava, C. N., Stone, E. M., and Sheffield, V. C. (1993a). Linkage of Bardet-Biedl syndrome to chromosome 16q and evidence for non- allelic genetic heterogeneity. Nat Genet *5*, 392-396.

Kwitek-Black, A. E., Carmi, R., Duyk, G. M., Buetow, K. H., Elbedour, K., Parvari, R., Yandava, C. N., Stone, E. M., and Sheffield, V. C. (1993b). Linkage of Bardet-Biedl syndrome to chromosome 16q and evidence for non-allelic genetic heterogeneity. Nat Genet *5*, 392-396.

Labrune, P., Fabre, M., Trioche, P., Estournet-Mathiaud, B., Grangeponte, M. C., Rambaud, C., Maurage, C., and Bernard, O. (1999). Jeune syndrome and liver disease: report of three cases treated with ursodeoxycholic acid. Am J Med Genet *87*, 324-328.

Lander, E. S., and Botstein, D. (1987). Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science *236*, 1567-1570.

Langer, L. O. (1968). Thoracic-pelvic-phalangeal dystrophy: asphyxiating thoracic dystrophy of the newborn, infantile thoracic dystrophy. Radiology *91*, 447-456.

Leader, B., and Leder, P. (2000). Formin-2, a novel formin homology protein of the cappuccino subfamily, is highly expressed in the developing and adult central nervous system. Mech Dev *93*, 221-231.

Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E., and de Graaff, E. (2003). A long-range *Shh* enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet *12*, 1725-1735.

Levitt, R., Dragwa, C., Hubbard, F., Scherbier-Heddema, T., Brown, D., Bonner, P., Rosen, S., Buetow, K., and Hudson, J. (1998). A chromosome specific high density screening set of tri- and tetranucleotide repeat polymorphisms. American Journal of Human Genetics *63*, A256.

Lew, D. J. (2002). Formin' actin filament bundles. Nat Cell Biol *4*, E29-30.

Li, Q. Y., Newbury-Ecob, R. A., Terrett, J. A., Wilson, D. I., Curtis, A. R., Yi, C. H., Gebuhr, T., Bullen, P. J., Robson, S. C., Strachan, T., *et al.* (1997). Holt-Oram syndrome is caused by mutations in *TBX5*, a member of the Brachyury (T) gene family. Nat Genet *15*, 21-29.

Li, X., Oghi, K. A., Zhang, J., Krones, A., Bush, K. T., Glass, C. K., Nigam, S. K., Aggarwal, A. K., Maas, R., Rose, D. W., and Rosenfeld, M. G. (2003). Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. Nature *426*, 247-254.

Liu, H. X., Cartegni, L., Zhang, M. Q., and Krainer, A. R. (2001). A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA1* and other genes. Nat Genet *27*, 55-58.

Maas, R., Elfering, S., Glaser, T., and Jepeal, L. (1994). Deficient outgrowth of the ureteric bud underlies the renal agenesis phenotype in mice manifesting the limb deformity (ld) mutation. Dev Dyn *199*, 214-228.

Majewski, E., Ozturk, B., and Gillessen-Kaesbach, G. (1996). Jeune syndrome with tongue lobulation and preaxial polydactyly, and Jeune syndrome with situs inversus and asplenia: compound heterozygosity Jeune-Mohr and Jeune-Ivemark? Am J Med Genet *63*, 74-79.

Martinez-Frias, M. L., Bermejo, E., Urioste, M., Egues, J., and Lopez Soler, J. A. (1993). Short rib-polydactyly syndrome (SRPS) with anencephaly and other central nervous system anomalies: a new type of SRPS or a more severe expression of a known SRPS entity? Am J Med Genet *47*, 782-787.

McGregor, L., Makela, V., Darling, S. M., Vrontou, S., Chalepakis, G., Roberts, C., Smart, N., Rutland, P., Prescott, N., Hopkins, J., *et al.* (2003). Fraser syndrome and mouse blebbed phenotype caused by mutations in *FRAS1/Fras1* encoding a putative extracellular matrix protein. Nat Genet *34*, 203-208.

Merino, R., Rodriguez-Leon, J., Macias, D., Ganan, Y., Economides, A. N., and Hurle, J. M. (1999). The BMP antagonist Gremlin regulates outgrowth, chondrogenesis and programmed cell death in the developing limb. Development *126*, 5515-5522.

Miano, M. G., Jacobson, S. G., Carothers, A., Hanson, I., Teague, P., Lovell, J., Cideciyan, A. V., Haider, N., Stone, E. M., Sheffield, V. C., and Wright, A. F. (2000). Pitfalls in homozygosity mapping. Am J Hum Genet *67*, 1348-1351.

Michos, O., Panman, L., Vintersten, K., Beier, K., Zeller, R., and Zuniga, A. (2004). Gremlin-mediated BMP antagonism induces the epithelial-mesenchymal feedback signaling controlling metanephric kidney and limb organogenesis. Development *131*, 3401-3410.

Ming, J. E., and Muenke, M. (2002). Multiple hits during early embryonic development: digenic diseases and holoprosencephaly. Am J Hum Genet *71*, 1017-1032.

Moon, A. M., Boulet, A. M., and Capecchi, M. R. (2000). Normal limb development in conditional mutants of Fgf4. Development *127*, 989-996.

Morgan, N. V., Bacchelli, C., Gissen, P., Morton, J., Ferrero, G. B., Silengo, M., Labrune, P., Casteels, I., Hall, C., Cox, P., *et al.* (2003). A locus for asphyxiating thoracic dystrophy, *ATD*, maps to chromosome 15q13. J Med Genet *40*, 431-435.

Mortlock, D. P., and Innis, J. W. (1997). Mutation of *HOXA13* in hand-foot-genital syndrome. Nat Genet *15*, 179-180.

Moudgil, A., Bagga, A., Kamil, E. S., Rimoin, D. L., Lachman, R. S., Cohen, A. H., and Jordan, S. C. (1998). Nephronophthisis associated with Ellis-van Creveld syndrome. Pediatr Nephrol *12*, 20-22.

Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B. R. (1996). Altered growth and branching patterns in synpolydactyly caused by mutations in *HOXD13*. Science *272*, 548-551.

Nagai, T., Nishimura, G., Kato, R., Hasegawa, T., Ohashi, H., and Fukushima, Y. (1995). Del(12)(p11.21p12.2) associated with an asphyxiating thoracic dystrophy or chondroectodermal dysplasia-like syndrome. Am J Med Genet *55*, 16-18.

Nezarati, M. M., and McLeod, D. R. (2002). Cenani-Lenz syndrome: report of a new case and review of the literature. Clin Dysmorphol *11*, 215-218.

Nishinakamura, R., Matsumoto, Y., Nakao, K., Nakamura, K., Sato, A., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Scully, S., Lacey, D. L., *et al.* (2001). Murine homolog of SALL1 is essential for ureteric bud invasion in kidney development. Development *128*, 3105-3115.

Oberklaid, F., Danks, D. M., Mayne, V., and Campbell, P. (1977). Asphyxiating thoracic dysplasia. Clinical, radiological, and pathological information on 10 patients. Arch Dis Child *52*, 758-765.

Ozcay, F., Derbent, M., Demirhan, B., Tokel, K., and Saatci, U. (2001). A family with Jeune syndrome. Pediatr Nephrol *16*, 623-626.

Pazour, G. J., Dickert, B. L., Vucica, Y., Seeley, E. S., Rosenbaum, J. L., Witman, G. B., and Cole, D. G. (2000). Chlamydomonas *IFT88* and its mouse homologue, polycystic kidney disease gene *tg737*, are required for assembly of cilia and flagella. J Cell Biol *151*, 709-718.

Pazour, G. J., and Rosenbaum, J. L. (2002). Intraflagellar transport and cilia-dependent diseases. Trends Cell Biol *12*, 551-555.

Pennica, D., Swanson, T. A., Welsh, J. W., Roy, M. A., Lawrence, D. A., Lee, J., Brush, J., Taneyhill, L. A., Deuel, B., Lew, M., *et al.* (1998). *WISP* genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors. Proc Natl Acad Sci U S A *95*, 14717-14722.

Perbal, B., Brigstock, D. R., and Lau, L. F. (2003). Report on the second international workshop on the *CCN* family of genes. Mol Pathol *56*, 80-85.

Percin, E. F., and Percin, S. (2003). Two unusual types of syndactyly in the same family; Cenani-Lenz type and "new" type versus severe type I syndactyly? Genet Couns *14*, 313-319.

Percin, E. F., Percin, S., Egilmez, H., Sezgin, I., Ozbas, F., and Akarsu, A. N. (1998). Mesoaxial complete syndactyly and synostosis with hypoplastic thumbs: an unusual combination or homozygous expression of syndactyly type I? J Med Genet *35*, 868-874.

Pfeiffer, R. A., and Meisel-Stosiek, M. (1982). Present nosology of the Cenani-Lenz type of syndactyly. Clin Genet *21*, 74-79.

Poggiani, C., Gasparoni, M. C., Mangili, G., and Colombo, A. (2000). Asphyxiating thoracic dysplasia in a lethal form: radiological and sonographic findings. Minerva Pediatr *52*, 63-67.

Pollak, M. R., Chou, Y. H., Cerda, J. J., Steinmann, B., La Du, B. N., Seidman, J. G., and Seidman, C. E. (1993). Homozygosity mapping of the gene for alkaptonuria to chromosome 3q2. Nat Genet *5*, 201-204.

Pring, M., Evangelista, M., Boone, C., Yang, C., and Zigmond, S. H. (2003). Mechanism of formin-induced nucleation of actin filaments. Biochemistry *42*, 486-496.

Qin, H., Diener, D. R., Geimer, S., Cole, D. G., and Rosenbaum, J. L. (2004). Intraflagellar transport (IFT) cargo: IFT transports flagellar precursors to the tip and turnover products to the cell body. J Cell Biol *164*, 255-266.

Rao, E., Blaschke, R. J., Marchini, A., Niesler, B., Burnett, M., and Rappold, G. A. (2001). The Leri-Weill and Turner syndrome homeobox gene *SHOX* encodes a cell-type specific transcriptional activator. Hum Mol Genet *10*, 3083-3091.

Richard, I., Broux, O., Hillaire, D., Cherif, D., Fougerousse, F., Cohen, D., and Beckmann, J. S. (1992). Mapping of the formin gene and exclusion as a candidate gene for the autosomal recessive form of limb-girdle muscular dystrophy. Hum Mol Genet *1*, 621-624.

Rosenbaum, J. L., and Witman, G. B. (2002). Intraflagellar transport. Nat Rev Mol Cell Biol *3*, 813-825.

Ruiz-Perez, V. L., Ide, S. E., Strom, T. M., Lorenz, B., Wilson, D., Woods, K., King, L., Francomano, C., Freisinger, P., Spranger, S., *et al.* (2000). Mutations in a new gene in Ellis-van Creveld syndrome and Weyers acrodental dysostosis. Nat Genet *24*, 283-286.

Ruiz-Perez, V. L., Tompson, S. W., Blair, H. J., Espinoza-Valdez, C., Lapunzina, P., Silva, E. O., Hamel, B., Gibbs, J. L., Young, I. D., Wright, M. J., and Goodship, J. A. (2003). Mutations in two nonhomologous genes in a head-to-head configuration cause Ellis-van Creveld syndrome. Am J Hum Genet *72*, 728-732.

Sagot, I., Klee, S. K., and Pellman, D. (2002). Yeast formins regulate cell polarity by controlling the assembly of actin cables. Nat Cell Biol *4*, 42-50.

Sanchez, M. P., Silos-Santiago, I., Frisen, J., He, B., Lira, S. A., and Barbacid, M. (1996). Renal agenesis and the absence of enteric neurons in mice lacking GDNF. Nature *382*, 70-73.

Sarafoglou, K., Funai, E. F., Fefferman, N., Zajac, L., Geneiser, N., Paidas, M. J., Greco, A., and Wallerstein, R. (1999). Short rib-polydactyly syndrome: more evidence of a continuous spectrum. Clin Genet *56*, 145-148.

Scholey, J. M. (2003). Intraflagellar transport. Annu Rev Cell Dev Biol *19*, 423-443.

Schuermann, M. J., Otto, E., Becker, A., Saar, K., Ruschendorf, F., Polak, B. C., Ala-Mello, S., Hoefele, J., Wiedensohler, A., Haller, M., *et al.* (2002). Mapping of gene loci for nephronophthisis type 4 and Senior-Loken syndrome, to chromosome 1p36. Am J Hum Genet *70*, 1240-1246.

Seven, M., Yuksel, A., Ozkilic, A., and Elcioglu, N. (2000). A variant of Cenani-Lenz type syndactyly. Genet Couns *11*, 41-47.

Shah, K. J. (1980). Renal lesion in Jeune's syndrome. Br J Radiol *53*, 432-436.

Shears, D. J., Guillen-Navarro, E., Sempere-Miralles, M., Domingo-Jimenez, R., Scambler, P. J., and Winter, R. M. (2002). Pseudodominant inheritance of Langer mesomelic dysplasia caused by a *SHOX* homeobox missense mutation. Am J Med Genet *110*, 153-157.

Shears, D. J., Vassal, H. J., Goodman, F. R., Palmer, R. W., Reardon, W., Superti-Furga, A., Scambler, P. J., and Winter, R. M. (1998). Mutation and deletion of the pseudoautosomal gene *SHOX* cause Leri-Weill dyschondrosteosis. Nat Genet *19*, 70-73.

Sheffield, V. C., Carmi, R., Kwitek-Black, A., Rokhlina, T., Nishimura, D., Duyk, G. M., Elbedour, K., Sunden, S. L., and Stone, E. M. (1994). Identification of a Bardet-

Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. Hum Mol Genet *3*, 1331-1335.

Sheffield, V. C., Nishimura, D. Y., and Stone, E. M. (1995). Novel approaches to linkage mapping. Curr Opin Genet Dev *5*, 335-341.

Shokeir, M. H., Houston, C. S., and Awen, C. F. (1971). Asphyxiating thoracic chondrodystrophy. Association with renal disease and evidence for possile heterozygous expression. J Med Genet *8*, 107-112.

Silengo, M., Gianino, P., Longo, P., Battistoni, G., and Defilippi, C. (2000). Dandy-Walker complex in a child with Jeune's asphyxiating thoracic dystrophy. Pediatr Radiol *30*, 430.

Singh, M., Ray, D., Paul, V. K., and Kumar, A. (1988). Hydrocephalus in asphyxiating thoracic dystrophy. Am J Med Genet *29*, 391-395.

Sun, X., Lewandoski, M., Meyers, E. N., Liu, Y. H., Maxson, R. E., Jr., and Martin, G. R. (2000). Conditional inactivation of Fgf4 reveals complexity of signalling during limb bud development. Nat Genet *25*, 83-86.

Sun, X., Mariani, F. V., and Martin, G. R. (2002). Functions of FGF signalling from the apical ectodermal ridge in limb development. Nature *418*, 501-508.

Takamiya, K., Kostourou, V., Adams, S., Jadeja, S., Chalepakis, G., Scambler, P. J., Huganir, R. L., and Adams, R. H. (2004). A direct functional link between the multi-PDZ domain protein GRIP1 and the Fraser syndrome protein Fras1. Nat Genet *36*, 172-177.

Takeda, H., Takami, M., Oguni, T., Tsuji, T., Yoneda, K., Sato, H., Ihara, N., Itoh, T., Kata, S. R., Mishina, Y., *et al.* (2002). Positional cloning of the gene *LIMBIN* responsible for bovine chondrodysplastic dwarfism. Proc Natl Acad Sci U S A *99*, 10549-10554.

Takigawa, M., Nakanishi, T., Kubota, S., and Nishida, T. (2003). Role of CTGF/HCS24/ecogenin in skeletal growth control. J Cell Physiol *194*, 256-266.

Tanaka, K. (2000). Formin family proteins in cytoskeletal control. Biochem Biophys Res Commun *267*, 479-481.

Temtamy, S. A., Ismail, S., and Nemat, A. (2003). Mild facial dysmorphism and quasidominant inheritance in Cenani-Lenz syndrome. Clin Dysmorphol *12*, 77-83.

Temtamy, S. A., and McKusick, V. A. (1978). The genetics of hand malformations. Birth Defects Orig Artic Ser *14*, 1-619.

Terwilliger, J. D., Ott, J. (1994). Handbook of human genetic linkage, John Hopkins).

Tickle, C. (2003). Patterning systems--from one end of the limb to the other. Dev Cell *4*, 449-458.

Tickle, C. (2004). The contribution of chicken embryology to the understanding of vertebrate limb development. Mech Dev *121*, 1019-1029.

Tongsong, T., Chanprapaph, P., and Thongpadungroj, T. (1999). Prenatal sonographic findings associated with asphyxiating thoracic dystrophy (Jeune syndrome). J Ultrasound Med *18*, 573-576.

Topol, L. Z., Modi, W. S., Koochekpour, S., and Blair, D. G. (2000). *DRM/GREMLIN* (*CKTSF1B1*) maps to human chromosome 15 and is highly expressed in adult and fetal brain. Cytogenet Cell Genet *89*, 79-84.

Torres, M., Gomez-Pardo, E., Dressler, G. R., and Gruss, P. (1995). Pax-2 controls multiple steps of urogenital development. Development *121*, 4057-4065.

Trumpp, A., Blundell, P. A., de la Pompa, J. L., and Zeller, R. (1992). The chicken limb deformity gene encodes nuclear proteins expressed in specific cell types during morphogenesis. Genes Dev *6*, 14-28.

Trupp, M., Arenas, E., Fainzilber, M., Nilsson, A. S., Sieber, B. A., Grigoriou, M., Kilkenny, C., Salazar-Grueso, E., Pachnis, V., and Arumae, U. (1996). Functional receptor for GDNF encoded by the c-ret proto-oncogene. Nature *381*, 785-789.

Turkel, S. B., Diehl, E. J., and Richmond, J. A. (1985). Necropsy findings in neonatal asphyxiating thoracic dystrophy. J Med Genet *22*, 112-118.

van Bokhoven, H., Celli, J., Kayserili, H., van Beusekom, E., Balci, S., Brussel, W., Skovby, F., Kerr, B., Percin, E. F., Akarsu, N., and Brunner, H. G. (2000). Mutation of the gene encoding the ROR2 tyrosine kinase causes autosomal recessive Robinow syndrome. Nat Genet *25*, 423-426.

Verma, I. C., Joseph, R., Bhargava, S., and Mehta, S. (1976). Split-hand and split-foot deformity inherited as an autosomal recessive trait. Clin Genet *9*, 8-14.

Vrontou, S., Petrou, P., Meyer, B. I., Galanopoulos, V. K., Imai, K., Yanagi, M., Chowdhury, K., Scambler, P. J., and Chalepakis, G. (2003). *Fras1* deficiency results in cryptophthalmos, renal agenesis and blebbed phenotype in mice. Nat Genet *34*, 209-214.

Wallar, B. J., and Alberts, A. S. (2003). The formins: active scaffolds that remodel the cytoskeleton. Trends Cell Biol *13*, 435-446.

Wang, C. C., Chan, D. C., and Leder, P. (1997). The mouse *formin* (*Fmn*) gene: genomic structure, novel exons, and genetic mapping. Genomics *39*, 303-311.

Wasserman, S. (1998). FH proteins as cytoskeletal organizers. Trends Cell Biol *8*, 111-115.

Watnick, T., and Germino, G. (2003). From cilia to cyst. Nat Genet *34*, 355-356.

Westendorf, J. J., and Koka, S. (2004). Identification of FHOD1-binding proteins and mechanisms of FHOD1-regulated actin dynamics. J Cell Biochem *92*, 29-41.

Whitley, C. B., Schwarzenberg, S. J., Burke, B. A., Freese, D. K., and Gorlin, R. J. (1987). Direct hyperbilirubinemia and hepatic fibrosis: a new presentation of Jeune syndrome (asphyxiating thoracic dystrophy). Am J Med Genet Suppl *3*, 211-220.

Wild, A., Kalff-Suske, M., Vortkamp, A., Bornholdt, D., Konig, R., and Grzeschik, K. H. (1997). Point mutations in human *GLI3* cause Greig syndrome. Hum Mol Genet *6*, 1979-1984.

Wilkie, A. O. (2003). Why study human limb malformations? J Anat *202*, 27-35.

Willems, P. J. (2000). Genetic causes of hearing loss. N Engl J Med *342*, 1101-1109.

Wilson, D. J., Weleber, R. G., and Beals, R. K. (1987). Retinal dystrophy in Jeune's syndrome. Arch Ophthalmol *105*, 651-657.

Winter, R. M. (1988). Malformation syndromes: a review of mouse/human homology. J Med Genet *25*, 480-487.

Winter, R. M., and Tickle, C. (1993). Syndactylies and polydactylies: embryological overview and suggested classification. Eur J Hum Genet *1*, 96-104.

Wong, M., Kireeva, M. L., Kolesnikova, T. V., and Lau, L. F. (1997). Cyr61, product of a growth factor-inducible immediate-early gene, regulates chondrogenesis in mouse limb bud mesenchymal cells. Dev Biol *192*, 492-508.

Woychik, R. P., Generoso, W. M., Russell, L. B., Cain, K. T., Cacheiro, N. L., Bultman, S. J., Selby, P. B., Dickinson, M. E., Hogan, B. L., and Rutledge, J. C. (1990a). Molecular and genetic characterization of a radiation-induced structural rearrangement in mouse chromosome 2 causing mutations at the limb deformity and agouti loci. Proc Natl Acad Sci U S A *87*, 2588-2592.

Woychik, R. P., Maas, R. L., Zeller, R., Vogt, T. F., and Leder, P. (1990b). 'Formins': proteins deduced from the alternative transcripts of the limb deformity gene. Nature *346*, 850-853.

Woychik, R. P., Stewart, T. A., Davis, L. G., D'Eustachio, P., and Leder, P. (1985). An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. Nature *318*, 36-40.

Wynshaw-Boris, A. (1996). Model mice and human disease. Nat Genet *13*, 259-260.

Wynshaw-Boris, A., Ryan, G., Deng, C. X., Chan, D. C., Jackson-Grusby, L., Larson, D., Dunmore, J. H., and Leder, P. (1997). The role of a single formin isoform in the limb and renal phenotypes of limb deformity. Mol Med *3*, 372-384.

Zeller, R., Haramis, A. G., Zuniga, A., McGuigan, C., Dono, R., Davidson, G., Chabanis, S., and Gibson, T. (1999). *Formin* defines a large family of morphoregulatory genes and functions in establishment of the polarising region. Cell Tissue Res *296*, 85-93.

Zeller, R., Jackson-Grusby, L., and Leder, P. (1989). The *limb deformity* gene is required for apical ectodermal ridge differentiation and anteroposterior limb pattern formation. Genes Dev *3*, 1481-1492.

Zigmond, S. H. (2004). Formin-induced nucleation of actin filaments. Curr Opin Cell Biol *16*, 99-105.

Zuniga, A., Haramis, A. P., McMahon, A. P., and Zeller, R. (1999). Signal relay by BMP antagonism controls the SHH/FGF4 feedback loop in vertebrate limb buds. Nature *401*, 598-602.

Zuniga, A., Michos, O., Spitz, F., Haramis, A. P., Panman, L., Galli, A., Vintersten, K., Klasen, C., Mansfield, W., Kuc, S., *et al.* (2004). Mouse *limb deformity* mutations disrupt a global control region within the large regulatory landscape required for *Gremlin* expression. Genes Dev *18*, 1553-1564.

Zuniga, A., and Zeller, R. (1999). *Gli3* (*Xt*) and *formin* (*ld*) participate in the positioning of the polarising region and control of posterior limb-bud identity. Development *126*, 13-21.

APPENDIX 1

Markers

Black (e.g. D8S1119) = 7 cM Single Chromosome Scan linkage marker set (Research Genetics, Inc.) – tri- and tetra-nucleotide repeats

Bold red (e.g. **D2S206**) = 10 cM ABI PRISM™ Linkage Mapping Set (Perkin Elmer Applied Biosystems Division) – di-nucleotide repeats

Italic blue (e.g. *D9S158*) = 10 cM Genome-wide Human Screening Set Version 8a (Research Genetics, Inc.) – tri- and tetra-nucleotide repeats

Bold purple (e.g. **D15S822**)= new markers chosen to fill in gaps – di-nucleotide repeats

Heterozygosity and distances

Het column gives heterozygosity of the marker available on the NCBI database uniSTS, Marshfield genetic map

Mb column gives distances in Mb from the top (tip of short arm) of the chromosome according to the UCSC Human Genome Assembly (April 2003 freeze) NCBI build 33

cM (Marshfield) column gives distances in cM from the top of the chromosome according to the Marshfield genetic map

cM (deCode) column gives distances in cM from the top of the chromosome according to the deCode genetic map

light blue fill indicates low heterozygosity

/ indicates heterozygosity or position unknown

Results

light green fill indicates individual is homozygous for that marker

green fill indicates two individuals from the same family are homozygous for 2 different alleles

bright green fill indicates candidate regions

light orange fill indicates result not clear

/ indicates no result could be obtained for that marker

Empty box indicates marker not tested

Key to genome-wide screen results tables.

Markers Chromosome 1	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5	Family 4 II.6
D1S243	0.86	0.06	0.00	/	243-255	243-255	240-262	240-262	252-254
D1S2845	0.84	3.91	8.85	6.43	230-236	224-230	224-230	224-228	240-242
D1S1612	0.83	7.71	16.22	13.82	113-125	113-113	109-113	113-121	118-126
ATA9B08	0.76	10.59	23.35	/	152-161	155-161	161-167	152-160	160-160
D1S1597	0.66	13.05	29.93	23.27	170-174	/	/	170-174	170-170
GATA29A05	0.71	16.73	37.05	/	182-182	182-202	186-198	186-202	186-202
D1S552	0.71	18.41	45.33	37.30	243-251	243-247	243-247	/	247-252
D1S1676	0.77	24.38	55.10	/	148-148	148-164	152-156	152-156	145-149
D1S1622	0.71	29.73	56.74	48.69	259-259	259-262	/	/	256-261
D1S2830	0.91	33.60	62.74	/	262-262	262-262	260-260	266-266	
GATA137F01	0.56	34.86	64.38	56.08	139-139	139-139	135-135	135-135	135-135
D1S255	0.74	37.07	65.47	58.66	78-78	72-78	72-72	72-78	
GATA129H04	0.88	41.19	72.59	65.87	217-245	217-217	221-233	229-233	213-225
D1S2874	0.73	47.26	75.66	70.35	231-231	235-235	232-238	230-234	232-236
D1S2134	0.83	47.63	75.66	71.29	266-266	266-266	266-266	266-266	257-291
D1S197	0.82	50.10	76.27	72.05	123-127	123-127	118-118	118-118	
D1S1150	/	57.52	85.68	81.30	333-361	333-357	349-353	353-353	342-342
GATA165C03	0.74	60.03	89.49	/	266-266	266-270	258-270	258-262	268-272
D1S209	0.80	61.33	93.86	87.51	164-164	164-164			
D1S1613	/	63.59	/	/	302-314	298-314	314-322	314-322	315-319
D1S2806	0.78	67.22	100.39	/	185-187	185-185			
D1S410	0.79	67.50	100.39	94.99	315-315	315-315	315-343	301-327	310-334
D1S1630	/	68.80	/	/	244-244	244-244	244-244	244-244	244-244
D1S219	0.82	69.22	101.48	97.02	154-160	154-154			160-168
D1S1665	0.74	73.61	102.02	99.62	226-226	226-230	226-234	226-230	226-226
D1S1728	0.67	12.26	109.04	/	157-165	157-165	157-157	157-165	158-162
D1S551	0.67	82.18	113.69	107.40	169-181	169-177	165-177	173-177	178-186
D1S2766	0.75	85.36	118.14	110.00	205-209	207-209	205-215	211-217	207-211
D1S2129	/	89.44	/	/	137-153	137-141	137-153	141-153	141-141

D1S1588	0.65	91.55	125.51	116.15	129-129	129-132	117-132	129-129	129-129
GATA45B07	0.67	99.13	131.34	/	263-263	263-263	263-267	271-271	272-272
D1S206	0.82	100.98	134.20	122.64	207-215	207-215			207-215
D1S1631	0.77	104.81	136.88	125.75	148-148	148-148	/	148-151	/
GATA176G01	0.84	106.81	140.39	/	198-202	198-202	194-198	190-194	179-191
D1S502	0.89	111.61	146.53	/	263-271	263-271			
D1S1191	/	113.86	147.60	135.51	123-123	123-123	119-127	123-127	123-123
D1S252	0.82	116.70	150.27	139.16	109-111	109-111	107-111	97-101	97-101
D1S534	0.83	118.82	151.88	140.68					201-213
D1S498	0.82	148.07	155.89	144.94	190-200	190-200	192-196	194-194	190-192
D1S1653	0.63	154.71	164.09	151.68	99-99	99-103	99-103	99-103	103-107
D1S1679	0.84	159.54	170.84	/	149-177	149-165	157-169	149-157	152-160
D1S1677	0.68	160.25	175.62	161.50	197-201	201-209	197-209	197-201	197-201
D1S1625	/	162.54	/	/	202-202	189-202	202-208	202-208	203-209
D1S196	0.73	164.30	181.49	169.40	308-308	308-308	308-308	308-324	308-318
D1S2851	0.83	167.01	188.32	172.90			273-277	/	270-274
D1S1589	0.77	170.84	192.05	176.25	200-212	200-200	203-203	203-203	205-205
D1S466	0.77	178.82	198.30	183.53			233-235	231-231	230-230
D1S518	0.84	184.07	202.19	188.02	197-205	193-205	197-201	201-213	206-206
D1S1660	0.79	195.07	212.44	194.98	230-242	230-230	242-250	234-242	239-243
D1S1678	0.67	200.26	218.46	203.75	/	/	303-315	299-303	299-303
D1S249	0.87	202.07	220.65	207.07	161-161	161-173	171-177	169-171	161-169
GATA124F08	0.66	/	/	/	237-237	237-237	237-237	237-241	235-239
D1S2141	0.82	211.81	233.38	219.23	244-248	240-244	248-260	232-260	248-256
D1S2871	0.84	218.68	241.26	/	229-241	230-230	232-234	232-232	234-236
D1S3462	0.77	228.42	247.23	235.37	259-259	259-259	253-259	259-259	263-266
D1S235	0.71	232.19	254.64	244.40	175-187	175-187	187-187	191-191	177-179
D1S2670	0.83	235.84	262.96	253.50	152-164	146-164	139-147	139-147	141-155
D1S547	0.77	238.05	267.51	259.16	290-290	290-306	286-290	286-290	294-294
D1S2811	0.89	239.89	274.53	/	159-159	149-159	142-144	/	146-148
D1S2836	0.78	242.96	285.75	271.84			267-267	267-267	271-271

Markers Chromosome 2	Het	Mb Build 33	cM Marshfield	cM deCode	Family 1 IV.1	Family 1 IV.3	Family 2 VI.3	Family 2 VI.5	Family 4 II.6
D2S2393	0.88	1.94	7.05	/	204-254	252-264			
D2S1780	/	3.30	/	/	/	/	/	/	315-315
D2S281	0.81	6.68	14.10	15.19	246-246	246-248?	247-251	247-251	
GATA116B01	0.77	8.10	17.88	19.64	203-203	181-203	185-185	185-185	186-198
D2S162	0.75	8.88	20.03	22.73	118-118	118-132	136-138	117-134	
D2S168	0.83	11.46	27.06	/	156-156	156-170			
D2S1400	0.66	11.62	27.60	29.24	118-118	118-118	114-118	114-114	114-114
D2S272	0.84	16.83	37.38	38.99			294-306	294-306	
D2S1360	0.86	17.47	38.33	39.96	174-178	178-178	178-178	178-178?	174-178
D2S305	0.72	19.39	38.87	/	316-328	328-328	324-324	324-324	
D2S2150	0.78	20.50	40.47	43.77			280-280	280-288	
D2S405	0.66	29.45	47.97	51.48	205-205	201-205	201-205	201-205	193-193
D2S1788	0.89	36.23	55.51	/	199-231	239-251	198-235	198-223	232-236
D2S1346	0.79	38.24	59.36	63.10	/	/	/	/	246-252
D2S2259	0.79	42.97	64.29	67.06	249-251	247-251	247-251	247-247	
D2S2739	0.90	49.59	73.61	/	317-321	321-321	297-325	/	308-324
D2S1337	/	59.34	/	/	174-178	178-178			
D2S441	0.76	68.21	86.82	91.23	134-142	134-134	138-150	138-150	138-150
D2S1774	/	77.18	/	/	152-152	152-152	148-152	144-152	152-172
D2S2733	0.71	78.49	99.41		/	/	/	/	173-176
D2S1777	0.65	78.50	99.41	102.50	195-199	/			
D2S139	0.82	79.73	101.56	103.88	107-127	107-127			
D2S1790	/	85.04	/	/	314-318	314-318	286-294	286-298	294-298
D2S113	0.79	96.76	111.21	/	206-206	206-206	222-222	206-222	
GATA176C01	0.77	102.19	114.42	116.13	/	/	/	/	227-227
D2S436	0.83	106.86	118.16	120.01	186-190	186-194	186-190	/	182-198
D2S160	0.78	113.09	122.96	124.94	210-210	206-210			
D2S410	0.83	116.33	125.18	127.38	163-167	169-173?	163-167	159-163	166-170
D2S1328	0.75	126.28	132.58	137.62	154-154	146-158	150-162	162-162	142-158
D2S1334	0.81	136.66	145.08	/	274-300	274-294	290-290	278-304	306-314

D2S442	0.66	137.47	147.40	149.66	202-202	198-202	198-206	202-206	198-198
D2S1399	0.80	148.42	152.04	/	138-138	163-163	134-146	138-146	138-138
D2S142	0.76	156.48	161.26	165.19	239-239	235-235			
D2S1353	0.80	159.76	164.51	167.91	147-147	153-153	150-156	/	144-153
D2S1776	0.72	169.84	173.00	/	302-302	298-302	298-298	/	298-298
D2S326	0.86	173.06	177.53	/			97-99	97-99	
GATA194A05	0.56	176.13	180.79	182.23	239-239	235-239	235-235	235-235	236-244
D2S364	0.80	182.99	186.21	187.67			248-248	248-248	
D2S1391	0.79	184.95	186.21	189.15	127-135	119-127	127-127	127-127	107-127
D2S117	0.82	195.58	194.45	194.63	201-205	191-205	200-204	200-204	
D2S1384	0.79	205.42	200.43	201.34	144-148	144-148	144-156	144-156	148-160
D2S155	0.77	207.15	202.92	/	204-204	204-204	204-206	204-206	
D2S2208	0.88	208.92	205.06	/	159-175	159-177			
D2S1649	/	/	208.00	/	111-131	119-131	111-127	111-127	111-115
D2S434	0.74	218.77	215.78	216.01	267-271	267-275	259-275	259-275	271-275
D2S339	0.77	222.90	221.13	224.87	150-152	150-154			
D2S2228	0.78	224.66	224.33	227.17			209-223	209-209	
D2S1363	0.79	227.23	227.00	229.15	178-186	/	170-170	170-170	170-186
D2S1370	0.78	229.73	231.00	/	245-249	249-249	249-257	249-257	243-251
D2S427	0.76	232.40	236.70	238.22	242-246	246-250	/	/	246-246
D2S206	0.79	233.91	240.79	240.03	142-142	142-150	146-146	148-148	
D2S1397	0.65	237.08	249.22	247.77	/	/	/	/	226-232
D2S338	0.79	237.52	250.54	248.37	283-289	283-285			
GATA178G09	0.72	238.36	251.94	249.60	181-181	181-181	/	181-185	169-189
D2S125	0.83	241.48	260.63	258.19	88-90	86-88			

Markers	Het	Mb	cM	cM	Family 1		Family 2		Family 4
Chromosome 3		Build 33	Marshfield	deCode	IV.1	IV.3	VI.3	VI.5	II.6
D3S2387	0.85	1.02	5.54	2.29	193-197	197-197	/	189-209	/
D3S1297	0.83	2.03	8.31	/			352-354	354-356	
D3S3050	0.76	3.28	14.46	10.31	205-213	205-205	205-205	205-205	208-212
D3S3030	/	6.25	/	/	111-131	111-131	127-127	127-127	123-127

Clin Genet 2001: 59: 203–205
Printed in Ireland. All rights reserved

CLINICAL GENETICS
ISSN 0009-9163

Letter to the Editor

Cenani–Lenz syndrome with renal hypoplasia is not linked to *FORMIN* or *GREMLIN*

SHORT REPORT

A locus for asphyxiating thoracic dystrophy, *ATD*, maps to chromosome 15q13

N V Morgan, C Bacchelli, P Gissen, J Morton, G B Ferrero, M Silengo, P Labrune, I Casteels, C Hall, P Cox, D A Kelly, R C Trembath, P J Scambler, E R Maher, F R Goodman, C A Johnson

J Med Genet 2003;**40**:431–435